(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 July 2001 (19.07.2001)

PCT

(10) International Publication Number
WO 01/51626 A2

(51) International Patent Classification[7]: C12N 15/00

(21) International Application Number: PCT/US01/01255

(22) International Filing Date: 9 January 2001 (09.01.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/480,879 10 January 2000 (10.01.2000) US

(71) Applicant: **ELIM BIOPHARMACEUTICALS, INC.** [US/US]; 901 Grandview Drive, South San Francisco, CA 94080 (US).

(72) Inventors: **LU, Xianbin**; 311 Faxon Avenue, San Francisco, CA 94112 (US). **SUN, Lihyun**; 140 Ware Road, Woodside, CA 94062 (US). **ZHANG, Yilin**; 140 Ware Road, Woodside, CA 94062 (US).

(74) Agents: **CHOW, Y., Ping** et al.; Heller Ehrman White & McAuliffe LLP, 275 Middlefield Road, Menlo Park, CA 94025-3506 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/51626 A2

(54) Title: NOVEL PLASMID DNA VECTORS

(57) Abstract: This invention relates to plasmid backbones and constructs wherein non-essential nucleotide sequences have been substantially removed. These backbones can be used to clone genes of interest. The constructs are useful in gene therapy, DNA vaccines and the *in vitro* production of polypeptides and/or proteins.

Novel Plasmid DNA Vectors

Field

The present invention relates generally to the field of plasmid constructs, specifically to constructs that are substantially free of non-essential sequences. The plasmids are useful for delivery of therapeutic nucleic acids to cells. In addition, this invention relates to methods of using those constructs as well as methods for selection of desired plasmids. Such constructs and methods are useful in most molecular biology applications, such as gene therapy, DNA vaccines and the *in vitro* production of polypeptides and/or proteins.

Background

Transgenesis refers to the process of introducing exogenous genes into the germ line of an organism. Transgenesis of humans would allow for the elimination of disease genes in a population of offspring, however, technical as well as ethical issues likely will prevent any transgenic experiments to be carried out with human eggs. Therefore, the ability to replace known disease genes with normal copies in afflicted humans is the ultimate goal of gene therapy. Human gene therapy protocols aim to introduce correcting copies of disease genes into somatic cells of the affected individual. Expression of a correct copy of an affected gene in somatic cells avoids transmission through the germ line, thereby, avoiding many of the ethical issues of transgenesis.

In the late 1980s, gene therapy was searching for a viable vector—a virus or lipid vehicle that would deliver DNA to cells in a way that would promote uptake and expression of the DNA. During the search it was discovered that injection of DNA plasmids in saline solution alone—that is, without the use of a vehicle—stimulated expression of foreign proteins in mice. The amounts of protein made by tissue receiving this so-called "naked DNA" were too low to correct an enzyme defect or provoke destruction of a tumor. However, the levels

were high enough to stimulate an immune response. DNA vaccination involves the deliberate introduction into cells of a DNA plasmid carrying an antigen-coding gene that induces an immune response.

The administration of plasmid DNA expression vectors causes intracellular synthesis of the encoded proteins and induction of long-lasting cellular and humoral immune responses. The exogenous DNA it must cross not only the cellular membrane but also the nuclear membrane since the transcription machinery is inside of the nucleus. The nuclear membrane has "pores," about 10 nm in diameter, through which molecules pass. For example, the nuclear proteins are synthesized in the cytoplasm and must enter the nucleus to transcribe the DNA to mRNA. Similarly, the mRNA must pass out of the nucleus to the cytoplasm to be translated. As indicated above, the nuclear pores are very small. For a given vector to cross membranes or pass through nuclear pores, a smaller plasmid would presumably encounter fewer obstacles during transfection. Thus, small sized vectors are desirable.

Plasmids are an essential element in genetic engineering and gene therapy. Plasmids are circular DNA molecules that can be introduced or transfected into bacterial cells by transformation which replicate autonomously in the cell. They offer several advantages as vectors, the most important of which is the ability to confer antibiotic resistance to the host cell. This allows direct selection for cells that receive and maintain recombinant DNA plasmids. Plasmids also allow the amplification of cloned DNA. Some plasmids are present in 20 to 50 copies during cell growth, and after the arrest of protein synthesis, as many as 1000 copies per cell of a plasmid can be generated.

Plasmid design and construction utilizes several key factors. First, plasmid replication origins determine plasmid copy number, which affects production yields. Plasmids that replicate to higher copy number can increase plasmid yield from a given volume of culture, but excessive copy number can be deleterious to the bacteria and lead to undesirable effects. Artificially constructed plasmids are

sometimes unstably maintained, leading to accumulation of plasmid-free cells and reduced production yields.

To overcome this, genes that code for antibiotic resistance phenotype are included on the plasmid. Antibiotics are often added to kill or inhibit plasmid-free cells. Most general-purpose cloning vectors contain ampicillin resistance (β-lactamase) genes while others have used the neo gene that encodes resistance to kanamycin and neomycin. Use of these genes can be problematic because of the potential for residual antibiotic in the purified DNA, which can cause an allergic reaction in a treated patient, as well as the potential for the transfer of antibiotic resistance genes to a potential pathogen.

Other plasmid elements that have been studied include partition elements. Such elements help stabilize plasmid maintenance independent of antibiotic selection. In addition, other elements promote monomerization of the plasmid. Some plasmids are prone to forming dimers, trimers and higher multimers that can reduce yield and interfere with maintenance, as well as generating a more complicated product profile. Multimer resolution elements have been employed to promote monomerization of plasmids.

In addition to elements that affect the behavior of the plasmid in the host bacteria, such as E. coli, plasmid vectors have also been shown to affect transfection and expression in eukaryotic cells. Certain plasmid sequences have been shown to reduce expression of eukaryotic genes in eukaryotic cells when carried in cis. Plasmid sequences also have been shown to fortuitously contain binding sites for transcriptional control proteins. This can cause inappropriate levels of gene expression in treated patients. In many cases, it is difficult or nearly impossible to predict when such unintended interactions will occur, unless empirical evaluation reveals the unexpected effects. Thus, a small plasmid that provides on the essential sequences is desirable.

Summary

The present invention is directed to novel plasmid DNA vectors that are substantially free of non-essential nucleotide sequences. Unnecessary or redundant DNA fragments are removed resulting in enhanced expression of the cloned gene in mammalian cells.

In one aspect of the invention there is a plasmid DNA vector consisting essentially of a replicon, and at least one other component selected from the group consisting of (i) promoter, (ii) intron, (iii) exogenous gene, (iv) transcription termination sequence, (v) selectable marker gene, and (vi) detectable marker gene wherein non-essential nucleotide sequences have been substantially removed.

In a second aspect of the invention there is provided a plasmid DNA vector consisting essentially of a promoter, origin of replication, polyadenylation signal, a bacterial resistance gene and a functional exogenous gene wherein the functional exogenous gene is operably linked to the promoter and wherein non-essential nucleotide sequences have been substantially removed.

In a further aspect of the invention there is provided a plasmid DNA vector consisting essentially of the following elements (i) a promoter functional in target cells, (ii) an intron, (iii) an exogenous gene, wherein said gene is operably linked to said promoter functional in target cells, (iv) a transcription termination sequence, (v) a replicon, and (vi) a selectable marker gene, wherein non-essential nucleotide sequences have been substantially removed.

In another aspect of the invention there is provided a plasmid DNA backbone consisting essentially of the following elements: (i) a promoter functional in target cells, (ii) an intron, (iii) a Multiple Cloning Site, wherein an exogenous gene may be inserted and operably linked to said promoter functional in target cells, (iv) a transcription termination sequence, (v) a replicon, and (vi) a selectable marker gene, wherein non-essential nucleotide sequences have been substantially removed.

In yet another aspect there is provided a host cell transfected with the inventive plasmid.

In a further aspect of the invention there is provided a method of transfecting a cell comprising contacting a cell with the inventive plasmid either *in vitro* or *in vivo* such that the plasmid enters the cell.

In another further aspect of the invention there is provided a method of gene therapy, either *in vivo* or *ex vivo*, comprising the administration to a patient in need of gene therapy a therapeutically effective amount of the inventive plasmid comprising an exogenous gene that is expressed in the patient.

Description of the Figures

Figure 1 is a representation of the novel plasmid DNA vector, peGT. Regions denoted G and H are sites for insertion of optional immunostimulatory sequences (ISS's). There may be greater than one, preferably two, ISS present.

Figure 2 is a representation of the novel plasmid DNA vector, peGT2.8, that does not contain the β-*galactosidase* gene. Regions denoted G and H are sites for insertion of optional immunostimulatory sequences (ISS's). There may be greater than one, preferably two, ISS present.

Figure 3 is a representation of the novel plasmid DNA vector, peGT2.6, that does not contain the SV40 intron nor the β-*galactosidase* gene. Regions denoted G and H are sites for insertion of optional immunostimulatory sequences (ISS's). There may be greater than one, preferably two, ISS present.

Figure 4 is a representation of the pVAX1/lacZ plasmid highlighting the regions where non-essential nucleotide sequences have been removed in the novel plasmids described herein. Regions denoted A (51 base pairs), B (114 base pairs) and C (112 base pairs) are non-essential nucleotide sequences which have been removed in the novel plasmids.

Figure 5 is a representation of the pCMVβ plasmid highlighting the regions where non-essential nucleotide sequences have been removed in the novel plasmids described herein. Regions denoted D (377 base pairs), E (157 base

pairs) and F (26 base pairs) are non-essential nucleotide sequences which have been removed in the novel plasmids.

Figure 6 is a graphical representation of Example 2. β-Gal expression in NIH3T3 cells by the novel vectors is compared with other commercially available vectors. β-Galactosidase activity of cells transfected with different plasmids using TransFast™ Reagent and grown in 6-well plates was assessed. Cell lysates were harvested 48 hours after transfection for β-galactosidase and protein assays. Expression of β-galactosidase was measured in milli-units of activity normalized by the amount of protein in each sample. The results represent the mean of triplicates; the error bars show the standard deviations. 1: non-transfected, 2&3: commercially available plasmids, 4: peGT, 5: peGT-ISS2, 6: peGT-ISS4.

Detailed Description

The invention will now be described in detail by way of reference only using the following definitions and examples. All patents and publications referred to herein are expressly incorporated by reference.

Definitions

As used herein, the following terms or abbreviations, whether used in the singular or plural, will have the meanings indicated:

The terms "in operable combination", "in operable order" and "operably linked" as used herein refer to the linkage of nucleic acid sequences in such a manner that a nucleic acid molecule capable of directing the transcription of a given gene and/or the synthesis of a desired protein molecule is produced. The term also refers to the linkage of amino acid sequences in such a manner so that a functional protein is produced.

A "detectable marker" gene is a gene that allows cells carrying the gene to be specifically detected (i.e., to be distinguished from cells that do not carry the marker gene). A large variety of such marker genes are known in the art. Preferred examples of such marker genes encode proteins appearing on cellular

surfaces, thereby facilitating simplified and rapid detection and/or cellular sorting. Also preferred are markers that allow visible detection of cells that carry the gene. Exemplary markers include, but are not limited to, a chemiluminescent, or enzymatic moieties.

A "selectable marker" gene is a gene that allows cells carrying the gene to be specifically selected for or against, in the presence of a corresponding selective agent. By way of illustration, an antibiotic resistance gene can be used as a positive selectable marker gene that allows a host cell to be positively selected for in the presence of the corresponding antibiotic. A variety of positive and negative selectable markers are known in the art, some of which are described herein. In a preferred embodiment the selectable marker is a bacterial resistance gene such as kanamycin, ampicillin, tetracycline, chloramphenicol and the like.

"Heterologous gene" means a nucleic acid molecule or gene sequence that is (1) normally found in the targeted cells, (2) normally found in targeted cells but not expressed at physiologically appropriate levels in targeted cells, (3) normally found in targeted cells but not expressed at optimal levels in certain pathological conditions, (4) not normally found in the targeted cells, (5) novel fragments of genes normally expressed or not expressed in targeted cells, (6) synthetic modifications of genes expressed or not expressed within targeted cells, (7) any other DNA which may be modified for expression in targeted cells and (8) any combination of the above. Usually the heterologous gene is a contiguous fragment of DNA. The genetic material that is introduced into targeted cells can be any DNA. The heterologous gene encodes an epitope of interest, a biological response modulator, a growth factor, a recognition sequence, a therapeutic gene, a fusion protein or combinations thereof. The term heterologous gene may be used interchangeably with transgene, foreign gene, or exogenous gene.

"Immunostimulatory sequence" (ISS) means a nucleotide sequence that enhances the immune response of a DNA vaccine. Preferred sequences include,

but are not limited to, the hexamers GACGTC, AGCGCT, and especially AACGTT.

"Intron" means a non-coding sequence of DNA within a gene, i.e., an intervening sequence in the gene. It is transcribed into hnRNA but is then removed by RNA splicing in the nucleus, leaving a mature mRNA that is then translated in the cytoplasm. Introns are essential for normal expression of several mammalian genes. The inclusion of introns in the transcribed region enhances expression by an unspecified post-transcriptional mechanism. Introns may also contain transcriptional enhancers and thus enhance expression via a transcriptional mechanism.

"Multiple Cloning Site" means a synthetic DNA sequence encoding a series of restriction endonuclease recognition sites. These sites are engineered for convenient cloning of DNA into a vector at a specific position.

"Non-essential nucleotide sequence" means non-functional DNA sequences in a vector. These sequences are not required for the expression of the cloned gene, stabilization of transcribed products or plasmid replication. Other terms that may be used interchangeably herein to describe these sequences include unnecessary, superfluous, expendable, unneeded, unwanted, dispensable, immaterial or redundant.

"Promoter" means a region containing specific DNA sequences to which RNA polymerase binds before initiating the transcription of DNA into RNA.

"Replicon" means a DNA molecule that is able to initiate its own replication. A replicon must have an origin of replication and usually also has the necessary regulatory information required for the proper initiation of DNA replication. Plasmid copy number is one of the more important factors affecting plasmid DNA yield. Copy number is determined primarily by the replicon. Plasmids derived from pBR322 contain the ColE1 origin of replication from pMB1. This origin of replication is tightly controlled, resulting in approximately 25 copies of the plasmid per bacterial cell (low copy number). Plasmids derived

from pUC contain a mutated version of the ColE1 origin of replication, which results in reduced replication control and approximately 200-700 plasmid copies per cell (high copy number).

The term "substantially" should be given its ordinary and accustomed meaning, i.e., considerable, ample, large. Therefore, the phrase "substantially removed" means that non-essential sequences have been removed to a considerable, ample or large extent.

"Targeted cells" means either cells maintained, or propagated, in vivo or in vitro and capable of being transformed by plasmids as discussed herein.

"Transcription termination signal" means nucleotide sequences that function to stop transcription.

The inventive plasmid may be described by reference to its component elements. In the present invention, specific elements can be used to provide functionalities, e.g., plasmid copy number in bacteria and the ability to select plasmid-bearing bacterial cells, to a plasmid for delivery of nucleic acid and thus provide functionalities within a transformed cell or animal containing such a plasmid. Various plasmid elements have a variety of effects that influence plasmid use as gene therapy vectors. This includes the relative position and orientation of plasmid elements. Those skilled in the art will recognize that specific portions of these regions or elements can be identified as containing the functional aspects providing the desirable property. Such elements can be readily defined using routine techniques or their equivalents.

Component elements of the plasmid include, but are not limited to, a promoter, detectable marker gene, selectable marker gene, transcription termination signal, replicon, and intron. The inventive plasmid may contain genetic sequences coding for elements to assist in the purification of an expressed protein, e.g., His tags. The inventive plasmid will minimally comprise a replicon. The other components can be added or removed, depending on the purpose for which the plasmid will be used. For example, the promoter can be changed or

even deleted. A plasmid without a promoter can be used for determining the promoter activity of either non-naturally occurring synthetic sequences or naturally occurring sequences that may be cloned in. Similarly, the plasmid may contain no heterologous gene. In a preferred example, the inventive vectors of SEQ ID Nos 4-9 do not contain any functional gene. These vectors are designed to allow cloning any desired gene(s) into the vector. Likewise, the intron or polyA tail may be omitted to allow the insertion of another intron or polyA tail sequence, respectively. In some cases even the selection gene can be deleted. For example, the plasmid can be co-transformed with another plasmid, which will possess a selectable marker gene. Other selection methods utilizing the host's existing selectivity may also be used.

Additionally, any component may have more than one copy present in the plasmid. For example, one plasmid can have more than one functional gene, or contain more than one selectable marker genes.

The components may be in any orientation, any combination, and any order. For example, the beta-Gal gene, in plasmids of SEQ. ID Nos 1-3, is oriented 5' to 3', while the kanamycin resistance gene is oriented 3' to 5'.

The regulatory elements necessary for gene expression of a DNA molecule include: a promoter, an initiation codon, a stop codon, and a polyadenylation signal. In addition, enhancers are often required for gene expression. It is necessary that these elements be operably linked to the exogenous gene sequence that encodes the desired proteins when these elements are not already associated with the exogenous gene sequence to be inserted and that the regulatory elements are operable in the cell in which they are transfected.

Initiation codons and stop codons are generally considered to be part of a nucleotide sequence that encodes the desired protein. However, it is necessary that these elements are functional in the cell in which the gene construct is transfected. The initiation and termination codons must be in frame with the coding sequence.

Promoters and polyadenylation signals used must be functional within the cells of the individual. Examples of promoters useful to practice the present invention may be either prokaryotic or eukaryotic and include, but are not limited to, promoters from Simian Virus 40 (SV40), Mouse Mammary Tumor Virus (MMTV) promoter, Human Immunodeficiency Virus (HIV) such as the HIV Long Terminal Repeat (LTR) promoter, Moloney virus, ALV, Cytomegalovirus (CMV) such as the CMV immediate early promoter, Epstein Barr Virus (EBV), Rous Sarcoma Virus (RSV) as well as promoters from human genes such as human Actin, human Myosin, human Hemoglobin, human muscle creatinine and human metalothionein. Preferred embodiments include, but are not limited to, the eukaryotic promoters EF-1a, SV40, RSV, and to the prokaryotic promoters T7 , T3, SP6, EM7 and lac. The promoter may optionally be replaced by a Multiple Cloning Site for detection of endogenous promoter activity of the inserted exogenous DNA fragment.

Polyadenylation signals which may be positioned at the 3' end of the transgene include, but are not limited to, both synthetic or natural signals such as those derived from bovine growth hormone (BGH) and SV40. Examples of polyadenylation signals useful to practice the present invention include, but are not limited to, SV40 polyadenylation signals and LTR polyadenylation signals. In particular, the SV40 polyadenylation signal is preferred.

The use of an intron positioned into a transcription unit, i.e., the exogenous gene, of the invention in order to activate the splicing apparatus and enhance message stability is within the scope of the invention.

The novel plasmid DNA vectors are constructed to substantially remove non-essential nucleotide sequences. With reference to pUC18 and two commercially available plasmid vectors, using the conventional numbering of the base pairs in the vectors, the inventive plasmids have the following characteristics:

(1) When compared to the commercially available plasmid pVAX1/LacZ (Invitrogen; SEQ ID NO. 10), the novel plasmid DNA vectors have removed the

non-coding region before the pUC origin of replication. The 51 base pairs corresponding to 5092-5143 of pVAX1/LacZ (equivalent to base pairs 449-500 of pUC18, which is the 3' non-coding region of the lacZ gene in pUC18) were removed. The novel plasmid DNA vectors have also removed the unnecessary ampicillin promoter before the pUC origin of replication corresponding to the base pairs 5144-5258 of pVAX1/LacZ (equivalent to base pairs 754-868 of pUC18). The pVAX1/LacZ vector is a kanamycin resistant vector; the ampicillin promoter is non-essential. Additionally, the unnecessary region before the CMV promoter of pVAX1/LacZ corresponding to base pairs 1-112 has been removed. See Figure 4.

(2) When compared to the commercially available plasmid pCMVβ (Clontech; SEQ ID NO. 11), the novel plasmid DNA vectors have removed the 377 base pairs corresponding to 4513-4890 of pCMVβ (the lacZ promoter of pUC18) and the 157 base pairs corresponding to 7007-7164 of pCMVβ (the lacZ 5' coding region in pUC18). pUC18 contains a simple version of the lacZ gene different from the large lacZ gene in other vectors. The small pUC18 lacZ is redundant in the pCMVβ vector. Therefore, any fragments belonging to it were removed. Additionally, the unnecessary region before the CMV promoter of pCMVβ corresponding to base pairs 1-26 has been removed. See Figure 5.

The inserted heterologous gene in a plasmid of the present invention, preferably encodes, but is not limited to, an expression product comprising: an epitope of interest, a biological response modulator, a growth factor, a recognition sequence, a therapeutic gene or a fusion protein.

Expression of the heterologous gene in a suitable host cell, e.g., prokaryotic or eukaryotic cells, allows for the production of large quantities of the expressed gene product, i.e. polypeptide or protein. A host cell strain for use with the inventive plasmids may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage or

folding) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins. Appropriate cells lines or host systems are well known in the art and can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to 3T3, CHO, VERO, BHK, HeLa, COS, MDCK, 293, WI38, etc. and selection of the appropriate cell line is within the ambit of one skilled in the art. Yeast systems are useful for expression and analysis of eukaryotic proteins. Insect cells are useful for recombinant protein expression and production.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express the heterologous gene may be engineered. Following the introduction of the novel plasmids by methods known in the art (for example $CaPO_4$ precipitation, electroporation and the like), engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the novel plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines that express the desired heterologous gene on the cell surface, and are particularly useful in screening analogs for the gene product of the heterologous gene. This method is also suitable for the production of secreted polypeptides that may encoded by heterologous genes.

The present inventive DNA vectors also find use in gene therapy, i.e., the transfer of selected genes into a host with the hope of ameliorating or curing a disease state. The inventive plasmids replace an absent or faulty gene with a

working gene, so that the body can make the correct enzyme or protein and consequently eliminate the root cause of the disease. The working gene is a heterologous gene obtained from an exogenous source.

Gene therapy for genetically caused diseases, such as cystic fibrosis or muscular dystrophy, is based on the discovery that such diseases are caused by particular genetic defects. A therapy for the genetic defect may be possible if the healthy gene is supplied to the afflicted organism in a sufficient amount. Gene therapy not only enables the treatment of genetically caused diseases, but is also suitable for the treatment of tumors, and is suited as a new form of inoculation against infectious diseases, such as hepatitis, influenza, and HIV, to give but a few examples.

The gene therapies according to the present invention comprise about 1 picogram to about 1000 milligrams of DNA. The therapy may be localized to a specific region or organ, or it may be administered systemically. In some preferred embodiments, the therapy contains about 1 nanograms to about 1000 micrograms of DNA. In some preferred embodiments, the therapy contains about 0.1 micrograms to about 500 micrograms of DNA. In some preferred embodiments, the therapy contains about 1 to about 350 micrograms of DNA. In some preferred embodiments, the therapy contains about 25 to about 250 micrograms of DNA. In some preferred embodiments, the therapy contains about 100 micrograms DNA.

Any of the methods for gene therapy available in the art can be used according to the present invention. For general reviews of the methods of gene therapy, see Goldspiel et al., 1993, *Clinical Pharmacy* 12:488-505; Wu and Wu, 1991, *Biotherapy* 3:87-95; Tolstoshev, 1993, *Ann. Rev. Pharmacol. Toxicol.* 32:573-596; Mulligan, 1993, *Science* 260:926-932; and Morgan and Anderson, 1993, *Ann. Rev. Biochem.* 62:191-217; May, 1993, *TIBTECH* 11(5):155-215). Methods commonly known in the art of recombinant DNA technology for selecting and cloning the heterologous gene are described in Ausubel et al. (eds.),

1993, Current Protocols in Molecular Biology, John Wiley & Sons, NY; and Kriegler, 1990, Gene Transfer and Expression, A Laboratory Manual, Stockton Press, N.Y.

Delivery of the nucleic acid into a patient may be either direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid-carrying vector, or indirect, in which case, cells are first transformed with the nucleic acid *in vitro*, then transplanted into the patient. These two approaches are known, respectively, as *in vivo* or *ex vivo* gene therapy.

The host cells which contain the coding sequence and which express the biologically active gene product may be identified by at least four general approaches; (a) DNA-DNA or DNA-RNA hybridization; (b) the presence or absence of "marker" gene functions; (c) assessing the level of transcription as measured by the expression of the heterologous gene mRNA transcripts in the host cell; and (d) detection of the gene product as measured by immunoassay or by its biological activity.

The inventive plasmids described herein may comprise DNA sequences coding for polypeptides that have useful therapeutic applications. These inventive plasmids may be naked in the sense that they are free from any delivery vehicle that can act to facilitate entry into the cell, for example, the polynucleotide sequences are free of viral sequences, particularly any viral particles which may carry genetic information. Similarly, they may be free from, or naked with respect to, any material which promotes transfection, such as liposomal formulations, charged lipids such as Lipofectin™ or precipitating agents such as $CaPO_4$. However, it is within the scope of the invention that the inventive plasmids may be used in conjunction with other agents such as, for example, those listed above to enhance DNA delivery and transfection.

The genetic vaccines according to the present invention comprise about 1 picogram to about 1000 milligrams of DNA and a maximal dose of 10 grams is contemplated. In some preferred embodiments, the vaccines contain about 10

picograms to about 800 milligrams of DNA. In some preferred embodiments, the vaccines contain about 0.1 to about 500 milligrams of DNA. In some preferred embodiments, the vaccines contain about 1 to about 350 milligrams of DNA. In some preferred embodiments, the vaccines contain about 25 to about 250 milligrams of DNA. In some preferred embodiments, the vaccines contain about 100 milligrams DNA. However, it is within the skill of art to appropriately adjust the amount of DNA to the response desired. Multiple doses may be required and/or administered.

The genetic vaccines according to the present invention are formulated according to the mode of administration to be used. One having ordinary skill in the art can readily formulate a genetic vaccine that comprises a novel plasmid described herein. In cases where intramuscular injection is the chosen mode of administration, an isotonic formulation is preferably used. Generally, additives for isotonicity can include sodium chloride, dextrose, mannitol, sorbitol and lactose. In some cases, isotonic solutions such as phosphate buffered saline are preferred. Stabilizers include gelatin and albumin. The pharmaceutical preparations according to the present invention are provided sterile and pyrogen free.

In some cases, the use of ISS's may be desirable to enhance the genetic vaccines. Currently, there are no ISS containing vectors commercially available. The present inventive plasmids may contain zero or multiple ISS's. The presence of the ISS's may be without adverse effect on the expression of inserted exogenous heterologous gene. ISS's are palindromic DNA sequences that induce production of interferon and enhance natural killer cell activity. However, it has been noted that the ISS's may interfere with gene expression. See Sato et al. Science (1996) 273:352-354. Surprisingly, the ISS's in the inventive plasmids have not been found to interfere with gene expression.

The following are examples of the present invention using various plasmid elements to design and construct plasmids for delivery of nucleic acid to various tissues. In particular, the following are specific examples of the construct peGT

and its derivatives. The utility of the plasmids of the present invention is noted by use in nonviral human gene therapy, genetic vaccines, cloning and expression of genes, and other standard molecular biology uses.

Examples

The following preparations and examples are given to enable those skilled in the art to more clearly understand and to practice the present invention. They should not be considered as limiting the scope of the invention, but merely as being illustrative and representative thereof.

Example 1

Plasmid Construction

A plasmid containing the lacZ gene was constructed. A schematic of the inventive plasmid constructed and utilized in subsequent experiments is shown in Figure 1, corresponding to SEQ ID NO 1. The plasmid contains the following elements: a promoter, a splice site, the lacZ gene, a transcription termination signal, a replicon, and bacterial resistance gene. In other embodiments, the bacterial resistance gene may be flanked with zero, one (SEQ ID NO 2) or two (SEQ ID NO 3) copies of an ISS on each side of the gene. The plasmid vectors used in the experiments described below were constructed using the CMV promoter, SV40 intron, the lacZ gene, the SV40 polyadenylation sequence, pUC origin of replication, and the kanamycin resistance gene with either zero, one or two copies of an ISS on each side. Plasmids without an ISS on each side of the kanamycin resistance gene are designated peGT. Plasmids with one ISS copy on each side of the kanamycin resistance gene are designated peGT-ISS2. Plasmids with two ISS copies on each side of the kanamycin resistance gene are designated peGT-ISS4.

Example 2

Plasmid Expression

Expression of the lacZ gene incorporated into three inventive plasmid vectors (peGT, peGT-ISS2 and peGT-ISS4) described in Example 1 was compared with two commercially available plasmids, pCMVβ (Clontech Laboratories, Palo Alto, CA 94303) and pVAX1/LacZ (Invitrogen Corporation, Carlsbad, CA 92008) in NIH3T3 cells. Surprisingly, the inventive plasmids, despite their smaller size, did not have a reduced level of expression, but rather attained a higher level of expression.

Cells were plated one day before transfection at the density of $2.5x10^5$ cells/well in 6-well plates. Cells were transfected using TransFast™ Reagent (Promega, Madison, WI). The total volume of medium, DNA, and TransFast™ Reagent (Promega, Madison, WI) to add per well for 6-well plates was 1 ml. On the day of transfection, the appropriate amount of serum-free medium (pre-warmed to 37°C) was added to a sterile tube. 2.5 μg of plasmid DNA/well were added to the medium and vortexed, and 3 μl of TransFast™ Reagent (1mM)/μg of DNA (1:1 charge ratio of TransFast™ Reagent to DNA) were then added and vortexed immediately. For each plasmid, 3 wells of 6-well plates were used. The DNA/TransFast™ Reagent/medium mixtures were incubated for 15 minutes at room temperature. Growth medium was removed from the cells, and the DNA/TransFast™ Reagent/medium mixtures were briefly vortexed and added to each well (1ml/well). Cells were returned immediately to the 37°C incubator for 1 hour. At the end of the incubation period, cells were gently overlayed with 2 ml of complete media and were returned to the 37°C incubator for 48 hours.

β-galactosidase activity was measured using β-galactosidase Enzyme Assay System (Promega, Madison, WI). Cells were harvested 48 hours after transfection. Cells from each well of 6-well plates were washed twice with PBS (phosphate-buffered saline, 137mM NaCl, 2.7mM KCl, 8.1mM Na_2HPO_4, 1.47mM KH_2PO_4), and lysed with Reporter Lysis Buffer (200μl for each well).

After 15-minute incubation at room temperature, cells lysate was scraped, and collected into a microcentrifuge tube on ice. Cell lysate was then vortexed and centrifuged at 14,000xg, and the supernatant was transferred to a fresh tube. Adequate amount of cell extracts were taken and mixed with Reporter Lysis Buffer to 50µl, and 50µl of Assay Buffer was added. The mixtures were incubated at 37°C for 30 minutes, and the reactions were stopped by 150 µl of 1M sodium carbonate. β-galactosidase activities were determined by the absorbance reading at 420nm in a 96-well plate.

Protein amounts were determined by BCA Protein Assay Kit (Pierce, Rockford, IL) using BSA (bovine serum albumin) as standards. β-galactosidase activity was normalized by protein quantity in each well.

The novel plasmids were efficiently transfected into and demonstrated a high level of gene expression in the host cells. The inventive plasmids demonstrated at least a two-fold increase in β-galactosidase activity per milligram protein when compared to commercially available plasmids. See Figure 6. Compared to the pCMVβ plasmid (Column 2, Figure 6), the inventive plasmid peGT expressed greater than two times more β-galactosidase activity. Similarly, when compared to the pVAX1 plasmid (Column 3, Figure 6), the inventive plasmid peGT expressed about ten times more β-galactosidase activity. Thus, despite their smaller size the inventive plasmid expressed higher levels of the desired gene product.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity and understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.

Claims

1. A plasmid DNA vector consisting essentially of a replicon, and at least one other component selected from the group consisting of:

(i) promoter,

(ii) intron,

(iii) exogenous gene,

(iv) transcription termination sequence,

(v) selectable marker gene,

(vi) detectable marker gene and

(vii) an immunostimulatory sequence

wherein non-essential nucleotide sequences have been substantially removed.

2. A plasmid according to Claim 1 consisting essentially of a replicon and an exogenous gene.

3. A plasmid according to Claim 1 consisting essentially of a replicon, an exogenous gene and a selectable marker gene.

4. A plasmid according to Claim 1 consisting essentially of a replicon, an exogenous gene, a selectable marker gene and a promoter.

5. A plasmid according to Claim 1 wherein the exogenous gene is a detectable marker gene.

6. A plasmid DNA vector consisting essentially of a promoter, origin of replication, polyadenylation signal, a bacterial resistance gene and a functional exogenous gene wherein the functional exogenous gene is operably linked to the

promoter and wherein non-essential nucleotide sequences have been substantially removed.

7. A plasmid DNA vector consisting essentially of the following elements:

(i) a promoter functional in target cells,

(ii) an intron,

(iii) an exogenous gene, wherein said gene is operably linked to said promoter functional in target cells,

(iv) a transcription termination sequence,

(v) a replicon, and

(vi) a selectable marker gene,

wherein non-essential nucleotide sequences have been substantially removed.

8. A plasmid according to Claim 7 wherein the exogenous gene encodes a detectable marker.

9. A plasmid according to Claim 7 wherein the selectable marker is a bacterial resistance gene.

10. A plasmid DNA backbone consisting essentially of the following elements:

(i) a promoter functional in target cells,

(ii) an intron,

(iii) a Multiple Cloning Site, wherein an exogenous gene may be inserted and operably linked to said promoter functional in target cells,

(iv) a transcription termination sequence,

(v) a replicon, and

(vi) a selectable marker gene,

wherein non-essential nucleotide sequences have been substantially removed.

11. A host cell transfected with the plasmid of Claim 1.

12. A method of transfecting a cell comprising contacting a cell with said plasmid of Claim 1 *in vitro*, such that said plasmid enters said cell.

13. A method of transfecting a cell comprising contacting a cell with said plasmid of Claim 1 *in vivo*, such that said plasmid enters said cell.

14. A method of gene therapy comprising the administration to a patient in need of gene therapy a therapeutically effective amount of the plasmid of Claim 1 comprising an exogenous gene that is expressed in said patient.

15. A method of *ex vivo* gene therapy comprising the administration to a patient in need of *ex vivo* gene therapy a therapeutically effective amount of transfected cells comprising the plasmid of Claim 1 comprising an exogenous gene that is expressed in said cells and reintroducing said cells to said patient.

SEQUENCE LISTING

<110> Zhang, Yilin
Lu, Xianbin
Sun, Lihyun

<120> Novel Plasmid Vectors

<130> Novel plasmid vectors

<140>
<141>

<160> 11

<170> PatentIn Ver. 2.1

<210> 1
<211> 5919
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:novel DNA plasmid

<400> 1
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 60
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 120
atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 180
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 300
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 360
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 420
ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 480
acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg 540
ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc tccggactct 600
agaggatccg gtactcgagg aactgaaaaa ccagaaagtt aactggtaag tttagtcttt 660
ttgtctttta tttcaggtcc cggatccggt ggtggtgcaa atcaaagaac tgctcctcag 720
tggatgttgc ctttacttct aggcctgtac ggaagtgtta cttctgctct aaaagctgcg 780
gaattgtacc cgcggccgca attcccgggg atcgaaagag cctgctaaag caaaaaagaa 840
gtcaccatgg atcccgtcgt tttacaacgt cgtgactggg aaaaccctgg cgttacccaa 900
cttaatcgcc ttgcagcaca tccccctttc gccagctggc gtaatagcga agaggcccgc 960
accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg aatggcgctt tgcctggttt 1020
ccggcaccag aagcggtgcc ggaaagctgg ctggagtgcg atcttcctga ggccgatact 1080
gtcgtcgtcc cctcaaactg gcagatgcac ggttacgatg cgcccatcta caccaacgta 1140
acctatccca ttacggtcaa tccgccgttt gttcccacgg agaatccgac gggttgttac 1200
tcgctcacat ttaatgttga tgaaagctgg ctacaggaag gccagacgcg aattattttt 1260

gatggcgtta actcggcgtt tcatctgtgg tgcaacgggc gctgggtcgg ttacggccag 1320
gacagtcgtt tgccgtctga atttgacctg agcgcatttt tacgcgccgg agaaaaccgc 1380
ctcgcggtga tggtgctgcg ttggagtgac ggcagttatc tggaagatca ggatatgtgg 1440
cggatgagcg gcattttccg tgacgtctcg ttgctgcata aaccgactac acaaatcagc 1500
gatttccatg ttgccactcg ctttaatgat gatttcagcc gcgctgtact ggaggctgaa 1560
gttcagatgt gcggcgagtt gcgtgactac ctacgggtaa cagtttcttt atggcagggt 1620
gaaacgcagg tcgccagcgg caccgcgcct ttcggcggtg aaattatcga tgagcgtggt 1680
ggttatgccg atcgcgtcac actacgtctg aacgtcgaaa acccgaaact gtggagcgcc 1740
gaaatcccga atctctatcg tgcggtggtt gaactgcaca ccgccgacgg cacgctgatt 1800
gaagcagaag cctgcgatgt cggtttccgc gaggtgcgga ttgaaaatgg tctgctgctg 1860
ctgaacggca agccgttgct gattcgaggc gttaaccgtc acgagcatca tcctctgcat 1920
ggtcaggtca tggatgagca gacgatggtg caggatatcc tgctgatgaa gcagaacaac 1980
tttaacgccg tgcgctgttc gcattatccg aaccatccgc tgtggtacac gctgtgcgac 2040
cgctacggcc tgtatgtggt ggatgaagcc aatattgaaa cccacggcat ggtgccaatg 2100
aatcgtctga ccgatgatcc gcgctggcta ccggcgatga gcgaacgcgt aacgcgaatg 2160
gtgcagcgcg atcgtaatca cccgagtgtg atcatctggt cgctggggaa tgaatcaggc 2220
cacggcgcta atcacgacgc gctgtatcgc tggatcaaat ctgtcgatcc ttcccgcccg 2280
gtgcagtatg aaggcggcgg agccgacacc acggccaccg atattatttg cccgatgtac 2340
gcgcgcgtgg atgaagacca gcccttcccg gctgtgccga aatggtccat caaaaaatgg 2400
ctttcgctac ctggagagac gcgcccgctg atcctttgcg aatacgccca cgcgatgggt 2460
aacagtcttg gcggtttcgc taaatactgg caggcgtttc gtcagtatcc ccgtttacag 2520
ggcggcttcg tctgggactg ggtggatcag tcgctgatta aatatgatga aaacggcaac 2580
ccgtggtcgg cttacggcgg tgattttggc gatacgccga acgatcgcca gttctgtatg 2640
aacggtctgg tctttgccga ccgcacgccg catccagcgc tgacggaagc aaaacaccag 2700
cagcagtttt tccagttccg tttatccggg caaaccatcg aagtgaccag cgaatacctg 2760
ttccgtcata gcgataacga gctcctgcac tggatggtgg cgctggatgg taagccgctg 2820
gcaagcggtg aagtgcctct ggatgtcgct ccacaaggta aacagttgat tgaactgcct 2880
gaactaccgc agccggagag cgccgggcaa ctctggctca cagtacgcgt agtgcaaccg 2940
aacgcgaccg catggtcaga agccgggcac atcagcgcct ggcagcagtg gcgtctggcg 3000
gaaaacctca gtgtgacgct ccccgccgcg tcccacgcca tcccgcatct gaccaccagc 3060
gaaatggatt tttgcatcga gctgggtaat aagcgttggc aatttaaccg ccagtcaggc 3120
tttctttcac agatgtggat tggcgataaa aaacaactgc tgacgccgct gcgcgatcag 3180
ttcacccgtg caccgctgga taacgacatt ggcgtaagtg aagcgacccg cattgaccct 3240
aacgcctggg tcgaacgctg gaaggcggcg ggccattacc aggccgaagc agcgttgttg 3300
cagtgcacgg cagatacact tgctgatgcg gtgctgatta cgaccgctca cgcgtggcag 3360
catcagggga aaaccttatt tatcagccgg aaaacctacc ggattgatgg tagtggtcaa 3420
atggcgatta ccgttgatgt tgaagtggcg agcgatacac cgcatccggc gcggattggc 3480
ctgaactgcc agctggcgca ggtagcagag cgggtaaact ggctcggatt agggccgcaa 3540
gaaaactatc ccgaccgcct tactgccgcc tgttttgacc gctgggatct gccattgtca 3600
gacatgtata ccccgtacgt cttcccgagc gaaaacggtc tgcgctgcgg gacgcgcgaa 3660
ttgaattatg gcccacacca gtggcgcggc gacttccagt tcaacatcag ccgctacagt 3720
caacagcaac tgatggaaac cagccatcgc catctgctgc acgcggaaga aggcacatgg 3780
ctgaatatcg acggtttcca tatggggatt ggtggcgacg actcctggag cccgtcagta 3840
tcggcggaat tacagctgag cgccggtcgc taccattacc agttggtctg gtgtcaaaaa 3900
taataatcta gtcgagaatt cgctagcaga catgataaga tacattgatg agtttggaca 3960
aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg atgctattgc 4020
tttatttgta accattataa gctgcaataa acaagttaac aacaacaatt gcattcattt 4080
tatgtttcag gttcaggggg aggtgtggga ggtttttttca tgtgagcaaa aggccagcaa 4140

aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgcccccct 4200
gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa 4260
agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg 4320
cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca 4380
cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa 4440
cccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaacccg 4500
gtaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg 4560
tatgtaggcg gtgctacaga gttcttgaag tggtggccta actacggcta cactagaagg 4620
acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc 4680
tcttgatccg gcaaacaaac caccgctggt agcggtggtt tttttgtttg caagcagcag 4740
attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac 4800
gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc 4860
ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag 4920
taaacttggt ctgacagtta gaagaactcg tcaagaaggc gatagaaggc gatgcgctgc 4980
gaatcgggag cggcgatacc gtaaagcacg aggaagcggt cagcccattc gccgccaagc 5040
tcttcagcaa tatcacgggt atccaacgct atgtcctgat agcggtccgc cacacccagc 5100
cggccacagt cgttgaatcc agaaaagcgg ccattttcca ccatgatatt cggcaagcag 5160
gcatcgccat ggatcacgac gagatcctcg ccgtcgggca tgcgcgcctt gagcctggcg 5220
aacagttcgg ctggcgcgag cccctgatgc tcttcgtcca gatcatcctg atcgacaaga 5280
ccggcttcca tccgagtacg tgctcgctcg atgcgatgtt tcgcttggtg gtcgaatggg 5340
caggtagccg gatcaagcgt atgcagccgc cgcattgcat cagccatgat ggatactttc 5400
tcggcaggag caaggcgaga tgacaggaga tcctgccccg gcacttcgcc caatagcagc 5460
cagtcccttc ccgcttcagt gacaacgtcg agcacagctg cgcaaggaac gcccgtcgtg 5520
gccagccacg atagccgcgc tgcctcgtcc tgcagttcat tcagggcacc ggacaggtcg 5580
gtcttgacaa aaagaaccgg gcgcccctgc gctgacagcc ggaacacggc ggcatcagag 5640
cagccgattg tctgttgtgc ccagtcatag ccgaatagcc tctccaccca agcggccgga 5700
gaacctgcgt gcaatccatc ttgttcaatc atgcgaaacg atcctcatcc tgtctcttga 5760
tcagatcttg atcccctgcg ccatcagatc cttggcggca agaaagccat ccagtttact 5820
ttgcagggct tcccaacctt accagagggc gccccagctg gcaattccgg ttcgcttgct 5880
gtccataaaa ccgcccagta gaaaataccg catcaggcg 5919

<210> 2
<211> 5931
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:novel DNA
plasmid

<400> 2
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 60
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 120
atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 180
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 300
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 360

atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 420
ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 480
acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg 540
ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc tccggactct 600
agaggatccg gtactcgagg aactgaaaaa ccagaaagtt aactggtaag tttagtcttt 660
ttgtctttta tttcaggtcc cggatccggt ggtggtgcaa atcaaagaac tgctcctcag 720
tggatgttgc ctttacttct aggcctgtac ggaagtgtta cttctgctct aaaagctgcg 780
gaattgtacc cgcggccgca attcccgggg atcgaaagag cctgctaaag caaaaaagaa 840
gtcaccatgg atcccgtcgt tttacaacgt cgtgactggg aaaaccctgg cgttacccaa 900
cttaatcgcc ttgcagcaca tccccctttc gccagctggc gtaatagcga agaggcccgc 960
accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg aatggcgctt tgcctggttt 1020
ccggcaccag aagcggtgcc ggaaagctgg ctggagtgcg atcttcctga ggccgatact 1080
gtcgtcgtcc cctcaaactg gcagatgcac ggttacgatg cgcccatcta caccaacgta 1140
acctatccca ttacggtcaa tccgccgttt gttcccacgg agaatccgac gggttgttac 1200
tcgctcacat ttaatgttga tgaaagctgg ctacaggaag gccagacgcg aattattttt 1260
gatggcgtta actcggcgtt tcatctgtgg tgcaacgggc gctgggtcgg ttacggccag 1320
gacagtcgtt tgccgtctga atttgacctg agcgcatttt tacgcgccgg agaaaaccgc 1380
ctcgcggtga tggtgctgcg ttggagtgac ggcagttatc tggaagatca ggatatgtgg 1440
cggatgagcg gcattttccg tgacgtctcg ttgctgcata aaccgactac acaaatcagc 1500
gatttccatg ttgccactcg ctttaatgat gatttcagcc gcgctgtact ggaggctgaa 1560
gttcagatgt gcggcgagtt gcgtgactac ctacgggtaa cagtttcttt atggcagggt 1620
gaaacgcagg tcgccagcgg caccgcgcct ttcggcggtg aaattatcga tgagcgtggt 1680
ggttatgccg atcgcgtcac actacgtctg aacgtcgaaa acccgaaact gtggagcgcc 1740
gaaatcccga atctctatcg tgcggtggtt gaactgcaca ccgccgacgg cacgctgatt 1800
gaagcagaag cctgcgatgt cggtttccgc gaggtgcgga ttgaaaatgg tctgctgctg 1860
ctgaacggca agccgttgct gattcgaggc gttaaccgtc acgagcatca tcctctgcat 1920
ggtcaggtca tggatgagca gacgatggtg caggatatcc tgctgatgaa gcagaacaac 1980
tttaacgccg tgcgctgttc gcattatccg aaccatccgc tgtggtacac gctgtgcgac 2040
cgctacggcc tgtatgtggt ggatgaagcc aatattgaaa cccacggcat ggtgccaatg 2100
aatcgtctga ccgatgatcc gcgctggcta ccggcgatga gcgaacgcgt aacgcgaatg 2160
gtgcagcgcg atcgtaatca cccgagtgtg atcatctggt cgctggggaa tgaatcaggc 2220
cacggcgcta atcacgacgc gctgtatcgc tggatcaaat ctgtcgatcc ttcccgcccg 2280
gtgcagtatg aaggcggcgg agccgacacc acggccaccg atattatttg cccgatgtac 2340
gcgcgcgtgg atgaagacca gcccttcccg gctgtgccga aatggtccat caaaaaatgg 2400
ctttcgctac ctggagagac gcgcccgctg atcctttgcg aatacgccca cgcgatgggt 2460
aacagtcttg gcggtttcgc taaatactgg caggcgtttc gtcagtatcc ccgtttacag 2520
ggcggcttcg tctgggactg ggtggatcag tcgctgatta aatatgatga aaacggcaac 2580
ccgtggtcgg cttacggcgg tgattttggc gatacgccga acgatcgcca gttctgtatg 2640
aacggtctgg tctttgccga ccgcacgccg catccagcgc tgacggaagc aaaacaccag 2700
cagcagtttt tccagttccg tttatccggg caaaccatcg aagtgaccag cgaatacctg 2760
ttccgtcata gcgataacga gctcctgcac tggatggtgg cgctggatgg taagccgctg 2820
gcaagcggtg aagtgcctct ggatgtcgct ccacaaggta aacagttgat tgaactgcct 2880
gaactaccgc agccggagag cgccgggcaa ctctggctca cagtacgcgt agtgcaaccg 2940
aacgcgaccg catggtcaga agccgggcac atcagcgcct ggcagcagtg gcgtctggcg 3000
gaaaacctca gtgtgacgct ccccgccgcg tcccacgcca tcccgcatct gaccaccagc 3060
gaaatggatt tttgcatcga gctgggtaat aagcgttggc aatttaaccg ccagtcaggc 3120
tttctttcac agatgtggat tggcgataaa aaacaactgc tgacgccgct gcgcgatcag 3180
ttcacccgtg caccgctgga taacgacatt ggcgtaagtg aagcgacccg cattgaccct 3240

aacgcctggg tcgaacgctg gaaggcggcg ggccattacc aggccgaagc agcgttgttg 3300
cagtgcacgg cagatacact tgctgatgcg gtgctgatta cgaccgctca cgcgtggcag 3360
catcagggga aaaccttatt tatcagccgg aaaacctacc ggattgatgg tagtggtcaa 3420
atggcgatta ccgttgatgt tgaagtggcg agcgatacac cgcatccggc gcggattggc 3480
ctgaactgcc agctggcgca ggtagcagag cgggtaaact ggctcggatt agggccgcaa 3540
gaaaactatc ccgaccgcct tactgccgcc tgttttgacc gctgggatct gccattgtca 3600
gacatgtata ccccgtacgt cttcccgagc gaaaacggtc tgcgctgcgg gacgcgcgaa 3660
ttgaattatg gcccacacca gtggcgcggc gacttccagt tcaacatcag ccgctacagt 3720
caacagcaac tgatggaaac cagccatcgc catctgctgc acgcggaaga aggcacatgg 3780
ctgaatatcg acggtttcca tatggggatt ggtggcgacg actcctggag cccgtcagta 3840
tcggcggaat tacagctgag cgccggtcgc taccattacc agttggtctg gtgtcaaaaa 3900
taataatcta gtcgagaatt cgctagcaga catgataaga tacattgatg agtttggaca 3960
aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg atgctattgc 4020
tttatttgta accattataa gctgcaataa acaagttaac aacaacaatt gcattcattt 4080
tatgtttcag gttcaggggg aggtgtggga ggtttttttca tgtgagcaaa aggccagcaa 4140
aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgcccccct 4200
gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa 4260
agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg 4320
cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca 4380
cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa 4440
ccccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaacccg 4500
gtaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg 4560
tatgtaggcg gtgctacaga gttcttgaag tggtggccta actacggcta cactagaagg 4620
acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc 4680
tcttgatccg gcaaacaaac caccgctggt agcggtggtt tttttgtttg caagcagcag 4740
attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac 4800
gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc 4860
ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag 4920
taaacttggt ctgacagtta acgttagaag aactcgtcaa gaaggcgata gaaggcgatg 4980
cgctgcgaat cgggagcggc gataccgtaa agcacgagga agcggtcagc ccattcgccg 5040
ccaagctctt cagcaatatc acgggtatcc aacgctatgt cctgatagcg gtccgccaca 5100
cccagccggc cacagtcgtt gaatccagaa aagcggccat tttccaccat gatattcggc 5160
aagcaggcat cgccatggat cacgacgaga tcctcgccgt cgggcatgcg cgccttgagc 5220
ctggcgaaca gttcggctgg cgcgagcccc tgatgctctt cgtccagatc atcctgatcg 5280
acaagaccgg cttccatccg agtacgtgct cgctcgatgc gatgtttcgc ttggtggtcg 5340
aatgggcagg tagccggatc aagcgtatgc agccgccgca ttgcatcagc catgatggat 5400
actttctcgg caggagcaag gcgagatgac aggagatcct gccccggcac ttcgcccaat 5460
agcagccagt cccttcccgc ttcagtgaca acgtcgagca cagctgcgca aggaacgccc 5520
gtcgtggcca gccacgatag ccgcgctgcc tcgtcctgca gttcattcag ggcaccggac 5580
aggtcggtct tgacaaaaag aaccgggcgc ccctgcgctg acagccggaa cacggcggca 5640
tcagagcagc cgattgtctg ttgtgcccag tcatagccga atagcctctc cacccaagcg 5700
gccggagaac ctgcgtgcaa tccatcttgt tcaatcatgc gaaacgatcc tcatcctgtc 5760
tcttgatcag atcttgatcc cctgcgccat cagatccttg gcggcaagaa agccatccag 5820
tttactttgc agggcttccc aaccttacca gagggcgccc cagctggcaa ttccggttcg 5880
cttgctgtcc ataaaaccgc ccagtaacgt tagaaaatac cgcatcaggc g 5931

<210> 3

<211> 5943
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:novel DNA plasmid

<400> 3
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 60
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 120
atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 180
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 300
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 360
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 420
ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 480
acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg 540
ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc tccggactct 600
agaggatccg gtactcgagg aactgaaaaa ccagaaagtt aactggtaag tttagtcttt 660
ttgtctttta tttcaggtcc cggatccggt ggtggtgcaa atcaaagaac tgctcctcag 720
tggatgttgc ctttacttct aggcctgtac ggaagtgtta cttctgctct aaaagctgcg 780
gaattgtacc cgcggccgca attcccgggg atcgaaagag cctgctaaag caaaaaagaa 840
gtcaccatgg atcccgtcgt tttacaacgt cgtgactggg aaaaccctgg cgttacccaa 900
cttaatcgcc ttgcagcaca tccccctttc gccagctggc gtaatagcga agaggcccgc 960
accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg aatggcgctt tgcctggttt 1020
ccggcaccag aagcggtgcc ggaaagctgg ctggagtgcg atcttcctga ggccgatact 1080
gtcgtcgtcc cctcaaactg gcagatgcac ggttacgatg cgcccatcta caccaacgta 1140
acctatccca ttacggtcaa tccgccgttt gttcccacgg agaatccgac gggttgttac 1200
tcgctcacat ttaatgttga tgaaagctgg ctacaggaag gccagacgcg aattattttt 1260
gatggcgtta actcggcgtt tcatctgtgg tgcaacgggc gctgggtcgg ttacggccag 1320
gacagtcgtt tgccgtctga atttgacctg agcgcatttt tacgcgccgg agaaaaccgc 1380
ctcgcggtga tggtgctgcg ttggagtgac ggcagttatc tggaagatca ggatatgtgg 1440
cggatgagcg gcattttccg tgacgtctcg ttgctgcata aaccgactac acaaatcagc 1500
gatttccatg ttgccactcg ctttaatgat gatttcagcc gcgctgtact ggaggctgaa 1560
gttcagatgt gcggcgagtt gcgtgactac ctacgggtaa cagtttcttt atggcagggt 1620
gaaacgcagg tcgccagcgg caccgcgcct ttcggcggtg aaattatcga tgagcgtggt 1680
ggttatgccg atcgcgtcac actacgtctg aacgtcgaaa acccgaaact gtggagcgcc 1740
gaaatcccga atctctatcg tgcggtggtt gaactgcaca ccgccgacgg cacgctgatt 1800
gaagcagaag cctgcgatgt cggtttccgc gaggtgcgga ttgaaaatgg tctgctgctg 1860
ctgaacggca agccgttgct gattcgaggc gttaaccgtc acgagcatca tcctctgcat 1920
ggtcaggtca tggatgagca gacgatggtg caggatatcc tgctgatgaa gcagaacaac 1980
tttaacgccg tgcgctgttc gcattatccg aaccatccgc tgtggtacac gctgtgcgac 2040
cgctacggcc tgtatgtggt ggatgaagcc aatattgaaa cccacggcat ggtgccaatg 2100
aatcgtctga ccgatgatcc gcgctggcta ccggcgatga gcgaacgcgt aacgcgaatg 2160
gtgcagcgcg atcgtaatca cccgagtgtg atcatctggt cgctggggaa tgaatcaggc 2220
cacggcgcta atcacgacgc gctgtatcgc tggatcaaat ctgtcgatcc ttcccgcccg 2280
gtgcagtatg aaggcggcgg agccgacacc acggccaccg atattatttg cccgatgtac 2340

gcgcgcgtgg atgaagacca gcccttcccg gctgtgccga aatggtccat caaaaaatgg 2400
ctttcgctac ctggagagac gcgcccgctg atcctttgcg aatacgccca cgcgatgggt 2460
aacagtcttg gcggtttcgc taaatactgg caggcgtttc gtcagtatcc ccgtttacag 2520
ggcggcttcg tctgggactg ggtggatcag tcgctgatta aatatgatga aaacggcaac 2580
ccgtggtcgg cttacggcgg tgattttggc gatacgccga acgatcgcca gttctgtatg 2640
aacggtctgg tctttgccga ccgcacgccg catccagcgc tgacggaagc aaaacaccag 2700
cagcagtttt tccagttccg tttatccggg caaaccatcg aagtgaccag cgaatacctg 2760
ttccgtcata gcgataacga gctcctgcac tggatggtgg cgctggatgg taagccgctg 2820
gcaagcggtg aagtgcctct ggatgtcgct ccacaaggta aacagttgat tgaactgcct 2880
gaactaccgc agccggagag cgccgggcaa ctctggctca cagtacgcgt agtgcaaccg 2940
aacgcgaccg catggtcaga agccgggcac atcagcgcct ggcagcagtg gcgtctggcg 3000
gaaaacctca gtgtgacgct ccccgccgcg tcccacgcca tcccgcatct gaccaccagc 3060
gaaatggatt tttgcatcga gctgggtaat aagcgttggc aatttaaccg ccagtcaggc 3120
tttctttcac agatgtggat tggcgataaa aaacaactgc tgacgccgct gcgcgatcag 3180
ttcacccgtg caccgctgga taacgacatt ggcgtaagtg aagcgacccg cattgaccct 3240
aacgcctggg tcgaacgctg gaaggcggcg ggccattacc aggccgaagc agcgttgttg 3300
cagtgcacgg cagatacact tgctgatgcg gtgctgatta cgaccgctca cgcgtggcag 3360
catcagggga aaaccttatt tatcagccgg aaaacctacc ggattgatgg tagtggtcaa 3420
atggcgatta ccgttgatgt tgaagtggcg agcgatacac cgcatccggc gcggattggc 3480
ctgaactgcc agctggcgca ggtagcagag cgggtaaact ggctcggatt agggccgcaa 3540
gaaaactatc ccgaccgcct tactgccgcc tgttttgacc gctgggatct gccattgtca 3600
gacatgtata ccccgtacgt cttcccgagc gaaaacggtc tgcgctgcgg gacgcgcgaa 3660
ttgaattatg gcccacacca gtggcgcggc gacttccagt tcaacatcag ccgctacagt 3720
caacagcaac tgatggaaac cagccatcgc catctgctgc acgcggaaga aggcacatgg 3780
ctgaatatcg acggtttcca tatggggatt ggtggcgacg actcctggag cccgtcagta 3840
tcggcggaat tacagctgag cgccggtcgc taccattacc agttggtctg gtgtcaaaaa 3900
taataatcta gtcgagaatt cgctagcaga catgataaga tacattgatg agtttggaca 3960
aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg atgctattgc 4020
tttatttgta accattataa gctgcaataa acaagttaac aacaacaatt gcattcattt 4080
tatgtttcag gttcaggggg aggtgtggga ggtttttttca tgtgagcaaa aggccagcaa 4140
aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgccccccct 4200
gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa 4260
agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg 4320
cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca 4380
cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa 4440
ccccccgttc agcccgaccg ctgcgccttа tccggtaact atcgtcttga gtccaacccg 4500
gtaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg 4560
tatgtaggcg gtgctacaga gttcttgaag tggtggccta actacggcta cactagaagg 4620
acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc 4680
tcttgatccg gcaaacaaac caccgctggt agcggtggtt tttttgtttg caagcagcag 4740
attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac 4800
gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc 4860
ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag 4920
taaacttggt ctgacagtta acgttaacgt tagaagaact cgtcaagaag gcgatagaag 4980
gcgatgcgct gcgaatcggg agcggcgata ccgtaaagca cgaggaagcg gtcagcccat 5040
tcgccgccaa gctcttcagc aatatcacgg gtatccaacg ctatgtcctg atagcggtcc 5100
gccacaccca gccggccaca gtcgttgaat ccagaaaagc ggccattttc caccatgata 5160
ttcggcaagc aggcatcgcc atggatcacg acgagatcct cgccgtcggg catgcgcgcc 5220

ttgagcctgg cgaacagttc ggctggcgcg agcccctgat gctcttcgtc cagatcatcc 5280
tgatcgacaa gaccggcttc catccgagta cgtgctcgct cgatgcgatg tttcgcttgg 5340
tggtcgaatg ggcaggtagc cggatcaagc gtatgcagcc gccgcattgc atcagccatg 5400
atggatactt tctcggcagg agcaaggcga gatgacagga gatcctgccc cggcacttcg 5460
cccaatagca gccagtccct tcccgcttca gtgacaacgt cgagcacagc tgcgcaagga 5520
acgcccgtcg tggccagcca cgatagccgc gctgcctcgt cctgcagttc attcagggca 5580
ccggacaggt cggtcttgac aaaaagaacc gggcgccgct gcgctgacag ccggaacacg 5640
gcggcatcag agcagccgat tgtctgttgt gcccagtcat agccgaatag cctctccacc 5700
caagcggccg gagaacctgc gtgcaatcca tcttgttcaa tcatgcgaaa cgatcctcat 5760
cctgtctctt gatcagatct tgatcccctg cgccatcaga tccttggcgg caagaaagcc 5820
atccagttta ctttgcaggg cttcccaacc ttaccagagg gcgccccagc tggcaattcc 5880
ggttcgcttg ctgtccataa aaccgcccag taacgttaac gttagaaaat accgcatcag 5940
gcg 5943

<210> 4
<211> 2624
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:novel DNA plasmid

<400> 4
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 60
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 120
atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 180
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 300
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 360
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 420
ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 480
acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg 540
ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc tccggactct 600
agactcgagt cgacaagctt gaattcgcta gcagacatga taagatacat tgatgagttt 660
ggacaaacca caactagaat gcagtgaaaa aaatgcttta tttgtgaaat ttgtgatgct 720
attgctttat ttgtaaccat tataagctgc aataaacaag ttaacaacaa caattgcatt 780
cattttatgt ttcaggttca gggggaggtg tgggaggttt tttcatgtga gcaaaaggcc 840
agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc 900
cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac 960
tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc 1020
tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata 1080
gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc 1140
acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 1200
acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 1260
cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta 1320
gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 1380

gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggtttttt gtttgcaagc 1440
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 1500
ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 1560
ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 1620
atgagtaaac ttggtctgac agttagaaga actcgtcaag aaggcgatag aaggcgatgc 1680
gctgcgaatc gggagcggcg ataccgtaaa gcacgaggaa gcggtcagcc cattcgccgc 1740
caagctcttc agcaatatca cgggtatcca acgctatgtc ctgatagcgg tccgccacac 1800
ccagccggcc acagtcgttg aatccagaaa agcggccatt ttccaccatg atattcggca 1860
agcaggcatc gccatggatc acgacgagat cctcgccgtc gggcatgcgc gccttgagcc 1920
tggcgaacag ttcggctggc gcgagccccc gatgctcttc gtccagatca tcctgatcga 1980
caagaccggc ttccatccga gtacgtgctc gctcgatgcg atgtttcgct tggtggtcga 2040
atgggcaggt agccggatca agcgtatgca gccgccgcat tgcatcagcc atgatggata 2100
ctttctcggc aggagcaagg cgagatgaca ggagatcctg ccccggcact tcgcccaata 2160
gcagccagtc ccttcccgct tcagtgacaa cgtcgagcac agctgcgcaa ggaacgcccg 2220
tcgtggccag ccacgatagc cgcgctgcct cgtcctgcag ttcattcagg gcaccggaca 2280
ggtcggtctt gacaaaaaga accgggcgcc cctgcgctga cagccggaac acggcggcat 2340
cagagcagcc gattgtctgt tgtgcccagt catagccgaa tagcctctcc acccaagcgg 2400
ccggagaacc tgcgtgcaat ccatcttgtt caatcatgcg aaacgatcct catcctgtct 2460
cttgatcaga tcttgatccc ctgcgccatc agatccttgg cggcaagaaa gccatccagt 2520
ttactttgca gggcttccca accttaccag agggcgcccc agctggcaat tccggttcgc 2580
ttgctgtcca taaaaccgcc cagtagaaaa taccgcatca ggcg 2624

<210> 5
<211> 2636
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:novel DNA
plasmid

<400> 5
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 60
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 120
atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 180
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 300
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 360
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 420
ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 480
acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg 540
ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc tccggactct 600
agactcgagt cgacaagctt gaattcgcta gcagacatga taagatacat tgatgagttt 660
ggacaaacca caactagaat gcagtgaaaa aaatgcttta tttgtgaaat ttgtgatgct 720
attgctttat ttgtaaccat tataagctgc aataaacaag ttaacaacaa caattgcatt 780
cattttatgt ttcaggttca gggggaggtg tgggaggttt tttcatgtga gcaaaaggcc 840
agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc 900

ccccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac 960
tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc 1020
tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata 1080
gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc 1140
acgaacccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 1200
acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 1260
cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta 1320
gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 1380
gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggtttttt gtttgcaagc 1440
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 1500
ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 1560
ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 1620
atgagtaaac ttggtctgac agttaacgtt agaagaactc gtcaagaagg cgatagaagg 1680
cgatgcgctg cgaatcggga gcggcgatac cgtaaagcac gaggaagcgg tcagcccatt 1740
cgccgccaag ctcttcagca atatcacggg tatccaacgc tatgtcctga tagcggtccg 1800
ccacacccag ccggccacag tcgttgaatc cagaaaagcg gccattttcc accatgatat 1860
tcggcaagca ggcatcgcca tggatcacga cgagatcctc gccgtcgggc atgcgcgcct 1920
tgagcctggc gaacagttcg gctggcgcga gcccctgatg ctcttcgtcc agatcatcct 1980
gatcgacaag accggcttcc atccgagtac gtgctcgctc gatgcgatgt ttcgcttggt 2040
ggtcgaatgg gcaggtagcc ggatcaagcg tatgcagccg ccgcattgca tcagccatga 2100
tggatacttt ctcggcagga gcaaggcgag atgacaggag atcctgcccc ggcacttcgc 2160
ccaatagcag ccagtccctt cccgcttcag tgacaacgtc gagcacagct gcgcaaggaa 2220
cgcccgtcgt ggccagccac gatagccgcg ctgcctcgtc ctgcagttca ttcagggcac 2280
cggacaggtc ggtcttgaca aaaagaaccg ggcgcccctg cgctgacagc cggaacacgg 2340
cggcatcaga gcagccgatt gtctgttgtg cccagtcata gccgaatagc ctctccaccc 2400
aagcggccgg agaacctgcg tgcaatccat cttgttcaat catgcgaaac gatcctcatc 2460
ctgtctcttg atcagatctt gatcccctgc gccatcagat ccttggcggc aagaaagcca 2520
tccagtttac tttgcagggc ttcccaacct taccagaggg cgccccagct ggcaattccg 2580
gttcgcttgc tgtccataaa accgcccagt aacgttagaa aataccgcat caggcg 2636

<210> 6
<211> 2648
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:novel DNA
plasmid

<400> 6
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 60
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 120
atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 180
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 300
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 360
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 420

ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 480
acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg 540
ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc tccggactct 600
agactcgagt cgacaagctt gaattcgcta gcagacatga taagatacat tgatgagttt 660
ggacaaacca caactagaat gcagtgaaaa aaatgcttta tttgtgaaat ttgtgatgct 720
attgctttat ttgtaaccat tataagctgc aataaacaag ttaacaacaa caattgcatt 780
cattttatgt ttcaggttca gggggaggtg tgggaggttt tttcatgtga gcaaaaggcc 840
agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc 900
cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac 960
tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc 1020
tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata 1080
gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc 1140
acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 1200
acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 1260
cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta 1320
gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 1380
gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggtttttt gtttgcaagc 1440
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 1500
ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 1560
ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 1620
atgagtaaac ttggtctgac agttaacgtt aacgttagaa gaactcgtca agaaggcgat 1680
agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg aagcggtcag 1740
cccattcgcc gccaagctct tcagcaatat cacgggtatc caacgctatg tcctgatagc 1800
ggtccgccac acccagccgg ccacagtcgt tgaatccaga aaagcggcca ttttccacca 1860
tgatattcgg caagcaggca tcgccatgga tcacgacgag atcctcgccg tcgggcatgc 1920
gcgccttgag cctggcgaac agttcggctg gcgcgagccc ctgatgctct tcgtccagat 1980
catcctgatc gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg 2040
cttggtggtc gaatgggcag gtagccggat caagcgtatg cagccgccgc attgcatcag 2100
ccatgatgga tactttctcg gcaggagcaa ggcgagatga caggagatcc tgccccggca 2160
cttcgcccaa tagcagccag tcccttcccg cttcagtgac aacgtcgagc acagctgcgc 2220
aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctcgtcctgc agttcattca 2280
gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga 2340
acacggcggc atcagagcag ccgattgtct gttgtgccca gtcatagccg aatagcctct 2400
ccacccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg cgaaacgatc 2460
ctcatcctgt ctcttgatca gatcttgatc ccctgcgcca tcagatcctt ggcggcaaga 2520
aagccatcca gtttactttg cagggcttcc caaccttacc agagggcgcc ccagctggca 2580
attccggttc gcttgctgtc cataaaaccg cccagtaacg ttaacgttag aaaataccgc 2640
atcaggcg 2648

<210> 7
<211> 2821
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:novel DNA plasmid

<400> 7

cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 60
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 120
atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 180
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 300
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 360
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 420
ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 480
acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg 540
ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc tccggactct 600
agaggatccg gtactcgagg aactgaaaaa ccagaaagtt aactggtaag tttagtcttt 660
ttgtctttta tttcaggtcc cggatccggt ggtggtgcaa atcaaagaac tgctcctcag 720
tggatgttgc ctttacttct aggcctgtac ggaagtgtta cttctgctct aaaagctgcg 780
gaattgtacc cgcggccgca ttaatactag tggtaccgaa ttcgctagca gacatgataa 840
gatacattga tgagtttgga caaaccacaa ctagaatgca gtgaaaaaaa tgctttattt 900
gtgaaatttg tgatgctatt gctttatttg taaccattat aagctgcaat aaacaagtta 960
acaacaacaa ttgcattcat tttatgtttc aggttcaggg ggaggtgtgg gaggtttttt 1020
catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt 1080
tttccatagg ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg 1140
gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg 1200
ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag 1260
cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc 1320
caagctgggc tgtgtgcacg aacccccccgt tcagcccgac cgctgcgcct tatccggtaa 1380
ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg 1440
taacaggatt agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc 1500
taactacggc tacactagaa ggacagtatt tggtatctgc gctctgctga agccagttac 1560
cttcggaaaa agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg 1620
tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt 1680
gatcttttct acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt 1740
catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa 1800
atcaatctaa agtatatatg agtaaacttg gtctgacagt tagaagaact cgtcaagaag 1860
gcgatagaag gcgatgcgct gcgaatcggg agcggcgata ccgtaaagca cgaggaagcg 1920
gtcagcccat tcgccgccaa gctcttcagc aatatcacgg gtatccaacg ctatgtcctg 1980
atagcggtcc gccacaccca gccggccaca gtcgttgaat ccagaaaagc ggccattttc 2040
caccatgata ttcggcaagc aggcatcgcc atggatcacg acgagatcct cgccgtcggg 2100
catgcgcgcc ttgagcctgg cgaacagttc ggctggcgcg agcccctgat gctcttcgtc 2160
cagatcatcc tgatcgacaa gaccggcttc catccgagta cgtgctcgct cgatgcgatg 2220
tttcgcttgg tggtcgaatg ggcaggtagc cggatcaagc gtatgcagcc gccgcattgc 2280
atcagccatg atggatactt tctcggcagg agcaaggcga gatgacagga gatcctgccc 2340
cggcacttcg cccaatagca gccagtccct tcccgcttca gtgacaacgt cgagcacagc 2400
tgcgcaagga acgcccgtcg tggccagcca cgatagccgc gctgcctcgt cctgcagttc 2460
attcagggca ccggacaggt cggtcttgac aaaaagaacc gggcgccccct gcgctgacag 2520
ccggaacacg gcggcatcag agcagccgat tgtctgttgt gcccagtcat agccgaatag 2580
cctctccacc caagcggccg gagaacctgc gtgcaatcca tcttgttcaa tcatgcgaaa 2640
cgatcctcat cctgtctctt gatcagatct tgatcccctg cgccatcaga tccttggcgg 2700
caagaaagcc atccagttta ctttgcaggg cttcccaacc ttaccagagg gcgccccagc 2760

tggcaattcc ggttcgcttg ctgtccataa aaccgcccag tagaaaatac cgcatcaggc 2820
g 2821

<210> 8
<211> 2833
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:novel DNA plasmid

<400> 8
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 60
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 120
atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 180
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 300
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 360
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 420
ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 480
acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg 540
ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc tccggactct 600
agaggatccg gtactcgagg aactgaaaaa ccagaaagtt aactggtaag tttagtcttt 660
ttgtctttta tttcaggtcc cggatccggt ggtggtgcaa atcaaagaac tgctcctcag 720
tggatgttgc ctttacttct aggcctgtac ggaagtgtta cttctgctct aaaagctgcg 780
gaattgtacc cgcggccgca ttaatactag tggtaccgaa ttcgctagca gacatgataa 840
gatacattga tgagtttgga caaaccacaa ctagaatgca gtgaaaaaaa tgctttattt 900
gtgaaatttg tgatgctatt gctttatttg taaccattat aagctgcaat aaacaagtta 960
acaacaacaa ttgcattcat tttatgtttc aggttcaggg ggaggtgtgg gaggtttttt 1020
catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt 1080
tttccatagg ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg 1140
gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg 1200
ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag 1260
cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc 1320
caagctgggc tgtgtgcacg aaccccccgt tcagcccgac cgctgcgcct tatccggtaa 1380
ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg 1440
taacaggatt agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc 1500
taactacggc tacactagaa ggacagtatt tggtatctgc gctctgctga agccagttac 1560
cttcggaaaa agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg 1620
tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt 1680
gatcttttct acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt 1740
catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa 1800
atcaatctaa agtatatatg agtaaacttg gtctgacagt taacgttaga agaactcgtc 1860
aagaaggcga tagaaggcga tgcgctgcga atcgggagcg gcgataccgt aaagcacgag 1920
gaagcggtca gcccattcgc cgccaagctc ttcagcaata tcacgggtat ccaacgctat 1980
gtcctgatag cggtccgcca cacccagccg gccacagtcg ttgaatccag aaaagcggcc 2040

attttccacc atgatattcg gcaagcaggc atcgccatgg atcacgacga gatcctcgcc 2100
gtcgggcatg cgcgccttga gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc 2160
ttcgtccaga tcatcctgat cgacaagacc ggcttccatc cgagtacgtg ctcgctcgat 2220
gcgatgtttc gcttggtggt cgaatgggca ggtagccgga tcaagcgtat gcagccgccg 2280
cattgcatca gccatgatgg atactttctc ggcaggagca aggcgagatg acaggagatc 2340
ctgccccggc acttcgccca atagcagcca gtcccttccc gcttcagtga caacgtcgag 2400
cacagctgcg caaggaacgc ccgtcgtggc cagccacgat agccgcgctg cctcgtcctg 2460
cagttcattc agggcaccgg acaggtcggt cttgacaaaa agaaccgggc gcccctgcgc 2520
tgacagccgg aacacggcgg catcagagca gccgattgtc tgttgtgccc agtcatagcc 2580
gaatagcctc tccacccaag cggccggaga acctgcgtgc aatccatctt gttcaatcat 2640
gcgaaacgat cctcatcctg tctcttgatc agatcttgat cccctgcgcc atcagatcct 2700
tggcggcaag aaagccatcc agtttacttt gcagggcttc ccaaccttac cagagggcgc 2760
cccagctggc aattccggtt cgcttgctgt ccataaaacc gcccagtaac gttagaaaat 2820
accgcatcag gcg 2833

<210> 9
<211> 2845
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:novel DNA plasmid

<400> 9
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 60
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 120
atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 180
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 300
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg 360
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg 420
ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt 480
acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg 540
ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc tccggactct 600
agaggatccg gtactcgagg aactgaaaaa ccagaaagtt aactggtaag tttagtcttt 660
ttgtctttta tttcaggtcc cggatccggt ggtggtgcaa atcaaagaac tgctcctcag 720
tggatgttgc ctttacttct aggcctgtac ggaagtgtta cttctgctct aaaagctgcg 780
gaattgtacc cgcggccgca ttaatactag tggtaccgaa ttcgctagca gacatgataa 840
gatacattga tgagtttgga caaaccacaa ctagaatgca gtgaaaaaaa tgctttattt 900
gtgaaatttg tgatgctatt gctttatttg taaccattat aagctgcaat aaacaagtta 960
acaacaacaa ttgcattcat tttatgtttc aggttcaggg ggaggtgtgg gaggtttttt 1020
catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt 1080
tttccatagg ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg 1140
gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg 1200
ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag 1260
cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc 1320

caagctgggc tgtgtgcacg aacccccccgt tcagcccgac cgctgcgcct tatccggtaa 1380
ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg 1440
taacaggatt agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc 1500
taactacggc tacactagaa ggacagtatt tggtatctgc gctctgctga agccagttac 1560
cttcggaaaa agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg 1620
ttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt 1680
gatcttttct acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt 1740
catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa 1800
atcaatctaa agtatatatg agtaaacttg gtctgacagt taacgttaac gttagaagaa 1860
ctcgtcaaga aggcgataga aggcgatgcg ctgcgaatcg ggagcggcga taccgtaaag 1920
cacgaggaag cggtcagccc attcgccgcc aagctcttca gcaatatcac gggtatccaa 1980
cgctatgtcc tgatagcggt ccgccacacc cagccggcca cagtcgttga atccagaaaa 2040
gcggccattt tccaccatga tattcggcaa gcaggcatcg ccatggatca cgacgagatc 2100
ctcgccgtcg ggcatgcgcg ccttgagcct ggcgaacagt tcggctggcg cgagcccctg 2160
atgctcttcg tccagatcat cctgatcgac aagaccggct tccatccgag tacgtgctcg 2220
ctcgatgcga tgtttcgctt ggtggtcgaa tgggcaggta gccggatcaa gcgtatgcag 2280
ccgccgcatt gcatcagcca tgatggatac tttctcggca ggagcaaggc gagatgacag 2340
gagatcctgc cccggcactt cgcccaatag cagccagtcc cttcccgctt cagtgacaac 2400
gtcgagcaca gctgcgcaag gaacgcccgt cgtggccagc cacgatagcc gcgctgcctc 2460
gtcctgcagt tcattcaggg caccggacag gtcggtcttg acaaaaagaa ccgggcgccc 2520
ctgcgctgac agccggaaca cggcggcatc agagcagccg attgtctgtt gtgcccagtc 2580
atagccgaat agcctctcca cccaagcggc cggagaacct gcgtgcaatc catcttgttc 2640
aatcatgcga aacgatcctc atcctgtctc ttgatcagat cttgatcccc tgcgccatca 2700
gatccttggc ggcaagaaag ccatccagtt tactttgcag ggcttcccaa ccttaccaga 2760
gggcgcccca gctggcaatt ccggttcgct tgctgtccat aaaaccgccc agtaacgtta 2820
acgttagaaa ataccgcatc aggcg 2845

<210> 10
<211> 6050
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:commercial plasmid

<400> 10
gactcttcgc gatgtacggg ccagatatac gcgttgacat tgattattga ctagttatta 60
atagtaatca attacggggt cattagttca tagcccatat atggagttcc gcgttacata 120
acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat tgacgtcaat 180
aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc aatgggtgga 240
ctatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc caagtacgcc 300
ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt acatgacctt 360
atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtgat 420
gcggttttgg cagtacatca atgggcgtgg atagcggttt gactcacggg gatttccaag 480
tctccacccc attgacgtca atgggagttt gttttggcac caaaatcaac gggactttcc 540
aaaatgtcgt aacaactccg ccccattgac gcaaatgggc ggtaggcgtg tacggtggga 600

ggtctatata agcagagctc tctggctaac tagagaaccc actgcttact ggcttatcga 660
aattaatacg actcactata gggagaccca agctggctag cgtttaaact taagcttggt 720
accgagctcg gatccactag tccagtgtgg tggaattctg cagatcgaaa cgatgataga 780
tcccgtcgtt ttacaacgtc gtgactggga aaaccctggc gttacccaac ttaatcgcct 840
tgcagcacat ccccctttcg ccagctggcg taatagcgaa gaggcccgca ccgatcgccc 900
ttcccaacag ttgcgcagcc tgaatggcga atggcgcttt gcctggtttc cggtaccaga 960
agcggtgccg gaaagctggc tggagtgcga tcttcctgag gccgatactg tcgtcgtccc 1020
ctcaaactgg cagatgcacg gttacgatgc gcccatctac accaacgtaa cctatcccat 1080
tacggtcaat ccgccgtttg ttcccacgga gaatccgacg ggttgttact cgctcacatt 1140
taatgttgat gaaagctggc tacaggaagg ccagacgcga attatttttg atggcgttaa 1200
ctcggcgttt catctgtggt gcaacgggcg ctgggtcggt tacggccagg acagtcgttt 1260
gccgtctgaa tttgacctga gcgcattttt acgcgccgga gaaaaccgcc tcgcggtgat 1320
ggtgctgcgt tggagtgacg gcagttatct ggaagatcag gatatgtggc ggatgagcgg 1380
cattttccgt gacgtctcgt tgctgcataa accgactaca caaatcagcg atttccatgt 1440
tgccactcgc tttaatgatg atttcagccg cgctgtactg gaggctgaag ttcagatgtg 1500
cggcgagttg cgtgactacc tacgggtaac agtttcttta tggcagggtg aaacgcaggt 1560
cgccagcggc accgcgcctt tcggcggtga aattatcgat gagcgtggtg gttatgccga 1620
tcgcgtcaca ctacgtctga acgtcgaaaa cccgaaactg tggagcgccg aaatcccgaa 1680
tctctatcgt gcggtggttg aactgcacac cgccgacggc acgctgattg aagcagaagc 1740
ctgcgatgtc ggtttccgcg aggtgcggat tgaaaatggt ctgctgctgc tgaacggcaa 1800
gccgttgctg attcgaggcg ttaaccgtca cgagcatcat cctctgcatg gtcaggtcat 1860
ggatgagcag acgatggtgc aggatatcct gctgatgaag cagaacaact ttaacgccgt 1920
gcgctgttcg cattatccga accatccgct gtggtacacg ctgtgcgacc gctacggcct 1980
gtatgtggtg gatgaagcca atattgaaac ccacggcatg gtgccaatga atcgtctgac 2040
cgatgatccg cgctggctac cggcgatgag cgaacgcgta acgcgaatgg tgcagcgcga 2100
tcgtaatcac ccgagtgtga tcatctggtc gctggggaat gaatcaggcc acggcgctaa 2160
tcacgacgcg ctgtatcgct ggatcaaatc tgtcgatcct tcccgcccgg tgcagtatga 2220
aggcggcgga gccgacacca cggccaccga tattatttgc ccgatgtacg cgcgcgtgga 2280
tgaagaccag cccttcccgg ctgtgccgaa atggtccatc aaaaaatggc tttcgctacc 2340
tggagagacg cgcccgctga tcctttgcga atacgcccac gcgatgggta acagtcttgg 2400
cggtttcgct aaatactggc aggcgtttcg tcagtatccc cgtttacagg gcggcttcgt 2460
ctgggactgg gtggatcagt cgctgattaa atatgatgaa aacggcaacc cgtggtcggc 2520
ttacggcggt gattttggcg atacgccgaa cgatcgccag ttctgtatga acggtctggt 2580
ctttgccgac cgcacgccgc atccagcgct gacggaagca aaacaccagc agcagttttt 2640
ccagttccgt ttatccgggc aaaccatcga agtgaccagc gaatacctgt tccgtcatag 2700
cgataacgag ctcctgcact ggatggtggc gctggatggt aagccgctgg caagcggtga 2760
agtgcctctg gatgtcgctc cacaaggtaa acagttgatt gaactgcctg aactaccgca 2820
gccggagagc gccgggcaac tctggctcac agtacgcgta gtgcaaccga acgcgaccgc 2880
atggtcagaa gccgggcaca tcagcgcctg gcagcagtgg cgtctggcgg aaaacctcag 2940
tgtgacgctc cccgccgcgt cccacgccat cccgcatctg accaccagcg aaatggattt 3000
ttgcatcgag ctgggtaata agcgttggca atttaaccgc cagtcaggct ttctttcaca 3060
gatgtggatt ggcgataaaa aacaactgct gacgccgctg cgcgatcagt tcacccgtgc 3120
accgctggat aacgacattg gcgtaagtga agcgacccgc attgacccta acgcctgggt 3180
cgaacgctgg aaggcggcgg gccattacca ggccgaagca gcgttgttgc agtgcacggc 3240
agatacactt gctgatgcgg tgctgattac gaccgctcac gcgtggcagc atcaggggaa 3300
aaccttattt atcagccgga aaacctaccg gattgatggt agtggtcaaa tggcgattac 3360
cgttgatgtt gaagtggcga gcgatacacc gcatccggcg cggattggcc tgaactgcca 3420
gctggcgcag gtagcagagc gggtaaactg gctcggatta gggccgcaag aaaactatcc 3480

cgaccgcctt actgccgcct gttttgaccg ctgggatctg ccattgtcag acatgtatac 3540
cccgtacgtc ttcccgagcg aaaacggtct gcgctgcggg acgcgcgaat tgaattatgg 3600
cccacaccag tggcgcggcg acttccagtt caacatcagc cgctacagtc aacagcaact 3660
gatggaaacc agccatcgcc atctgctgca cgcggaagaa ggcacatggc tgaatatcga 3720
cggtttccat atggggattg gtggcgacga ctcctggagc ccgtcagtat cggcggaatt 3780
ccagctgagc gccggtcgct accattacca gttggtctgg tgtcaaaaag cggccgctcg 3840
agtctagagg gcccgtttaa acccgctgat cagcctcgac tgtgccttct agttgccagc 3900
catctgttgt ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc actcccactg 3960
tcctttccta ataaaatgag gaaattgcat cgcattgtct gagtaggtgt cattctattc 4020
tgggggggtgg ggtggggcag gacagcaagg gggaggattg ggaagacaat agcaggcatg 4080
ctggggatgc ggtgggctct atggcttcta ctgggcggtt ttatggacag caagcgaacc 4140
ggaattgcca gctggggcgc cctctggtaa ggttgggaag ccctgcaaag taaactggat 4200
ggctttctcg ccgccaagga tctgatggcg caggggatca agctctgatc aagagacagg 4260
atgaggatcg tttcgcatga ttgaacaaga tggattgcac gcaggttctc cggccgcttg 4320
ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct ctgatgccgc 4380
cgtgttccgg ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg acctgtccgg 4440
tgccctgaat gaactgcaag acgaggcagc gcggctatcg tggctggcca cgacgggcgt 4500
tccttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc tgctattggg 4560
cgaagtgccg gggcaggatc tcctgtcatc tcaccttgct cctgccgaga aagtatccat 4620
catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc cattcgacca 4680
ccaagcgaaa catcgcatcg agcgagcacg tactcggatg gaagccggtc ttgtcgatca 4740
ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg ccaggctcaa 4800
ggcgagcatg cccgacggcg aggatctcgt cgtgacccat ggcgatgcct gcttgccgaa 4860
tatcatggtg gaaaatggcc gcttttctgg attcatcgac tgtggccggc tgggtgtggc 4920
ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc ttggcggcga 4980
atgggctgac cgcttcctcg tgctttacgg tatcgccgct cccgattcgc agcgcatcgc 5040
cttctatcgc cttcttgacg agttcttctg aattattaac gcttacaatt tcctgatgcg 5100
gtattttctc cttacgcatc tgtgcggtat ttcacaccgc atacaggtgg cacttttcgg 5160
ggaaatgtgc gcggaacccc tatttgttta tttttctaaa tacattcaaa tatgtatccg 5220
ctcatgagac aataaccctg ataaatgctt caataatagc acgtgctaaa acttcatttt 5280
taatttaaaa ggatctaggt gaagatcctt tttgataatc tcatgaccaa aatcccttaa 5340
cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga 5400
gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa aaaaaccacc gctaccagcg 5460
gtggtttgtt tgccggatca agagctacca actcttttte cgaaggtaac tggcttcagc 5520
agagcgcaga taccaaatac tgtccttcta gtgtagccgt agttaggcca ccacttcaag 5580
aactctgtag caccgcctac atacctcgct ctgctaatcc tgttaccagt ggctgctgcc 5640
agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg 5700
cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac 5760
accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga 5820
aaggcggaca ggtatccggt aagcggcagg gtcggaacag gagagcgcac gagggagctt 5880
ccagggggaa acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag 5940
cgtcgatttt tgtgatgctc gtcaggggggg cggagcctat ggaaaaacgc cagcaacgcg 6000
gccttttttac ggttcctggg cttttgctgg ccttttgctc acatgttctt 6050

<210> 11
<211> 7164
<212> DNA

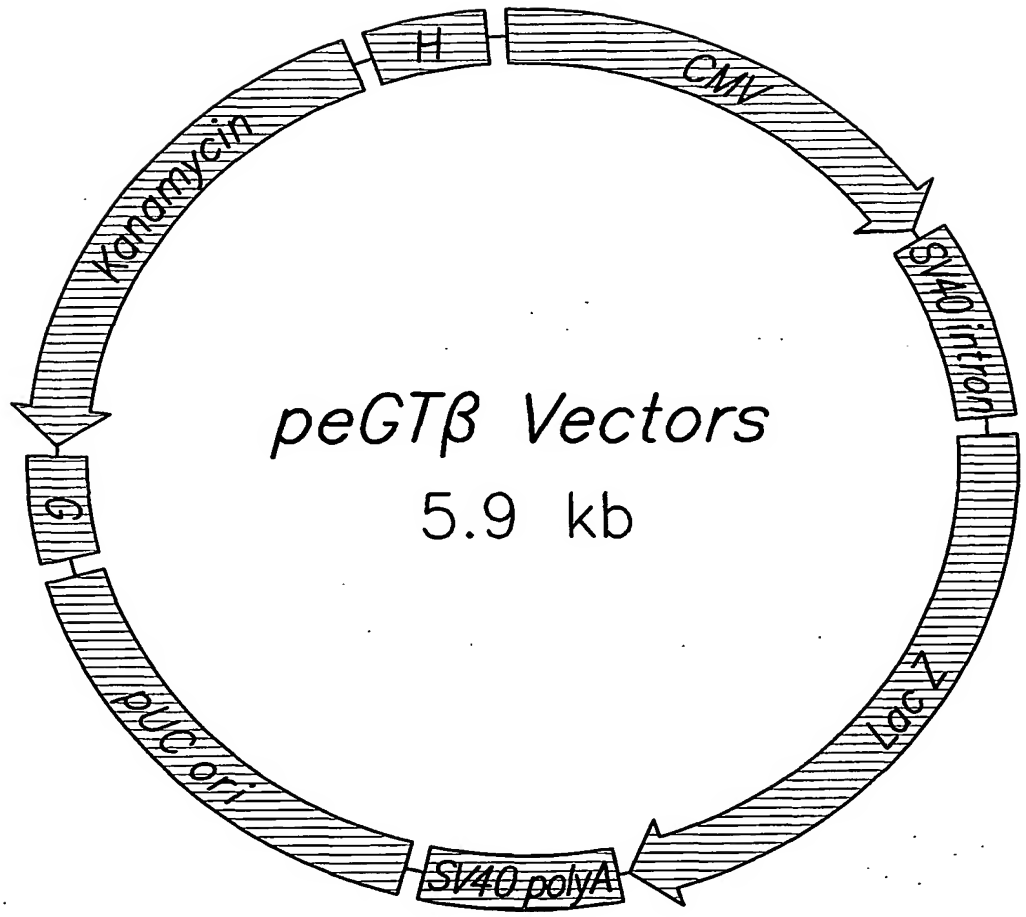
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:commercial plasmid

<400> 11
gaattcgagc ttgcatgcct gcaggtcgtt acataactta cggtaaatgg cccgcctggc 60
tgaccgccca acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg 120
ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg 180
gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa 240
tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac 300
atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta catcaatggg 360
cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga cgtcaatggg 420
agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa ctccgcccca 480
ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag agctcgttta 540
gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca tagaagacac 600
cgggaccgat ccagcctccg gactctagag gatccggtac tcgaggaact gaaaaaccag 660
aaagttaact ggtaagttta gtctttttgt cttttatttc aggtcccgga tccggtggtg 720
gtgcaaatca aagaactgct cctcagtgga tgttgccttt acttctaggc ctgtacggaa 780
gtgttacttc tgctctaaaa gctgcggaat tgtacccgcg gccgcaattc ccggggatcg 840
aaagagcctg ctaaagcaaa aaagaagtca ccatgtcgtt tactttgacc aacaagaacg 900
tgattttcgt tgccggtctg ggaggcattg gtctggacac cagcaaggag ctgctcaagc 960
gcgatcccgt cgttttacaa cgtcgtgact gggaaaaccc tggcgttacc caacttaatc 1020
gccttgcagc acatcccccc ttcgccagct ggcgtaatag cgaagaggcc cgcaccgatc 1080
gcccttccca acagttgcgc agcctgaatg gcgaatggcg ctttgcctgg tttccggcac 1140
cagaagcggt gccggaaagc tggctggagt gcgatcttcc tgaggccgat actgtcgtcg 1200
tcccctcaaa ctggcagatg cacggttacg atgcgcccat ctacaccaac gtaacctatc 1260
ccattacggt caatccgccg tttgttccca cggagaatcc gacgggttgt tactcgctca 1320
catttaatgt tgatgaaagc tggctacagg aaggccagac gcgaattatt tttgatggcg 1380
ttaactcggc gtttcatctg tggtgcaacg ggcgctgggt cggttacggc caggacagtc 1440
gtttgccgtc tgaatttgac ctgagcgcat ttttacgcgc cggagaaaac cgcctcgcgg 1500
tgatggtgct gcgttggagt gacggcagtt atctggaaga tcaggatatg tggcggatga 1560
gcggcatttt ccgtgacgtc tcgttgctgc ataaaccgac tacacaaatc agcgatttcc 1620
atgttgccac tcgctttaat gatgatttca gccgcgctgt actggaggct gaagttcaga 1680
tgtgcggcga gttgcgtgac taccttacggg taacagtttc tttatggcag ggtgaaacgc 1740
aggtcgccag cggcaccgcg cctttcggcg gtgaaattat cgatgagcgt ggtggttatg 1800
ccgatcgcgt cacactacgt ctgaacgtcg aaaacccgaa actgtggagc gccgaaatcc 1860
cgaatctcta tcgtgcggtg gttgaactgc acaccgccga cggcacgctg attgaagcag 1920
aagcctgcga tgtcggtttc cgcgaggtgc ggattgaaaa tggtctgctg ctgctgaacg 1980
gcaagccgtt gctgattcga ggcgttaacc gtcacgagca tcatcctctg catggtcagg 2040
tcatggatga gcagacgatg gtgcaggata tcctgctgat gaagcagaac aactttaacg 2100
ccgtgcgctg ttcgcattat ccgaaccatc cgctgtggta cacgctgtgc gaccgctacg 2160
gcctgtatgt ggtggatgaa gccaatattg aaacccacgg catggtgcca atgaatcgtc 2220
tgaccgatga tccgcgctgg ctaccggcga tgagcgaacg cgtaacgcga atggtgcagc 2280
gcgatcgtaa tcacccgagt gtgatcatct ggtcgctggg gaatgaatca ggccacggcg 2340
ctaatcacga cgcgctgtat cgctggatca aatctgtcga tccttcccgc ccggtgcagt 2400
atgaaggcgg cggagccgac accacggcca ccgatattat ttgcccgatg tacgcgcgcg 2460

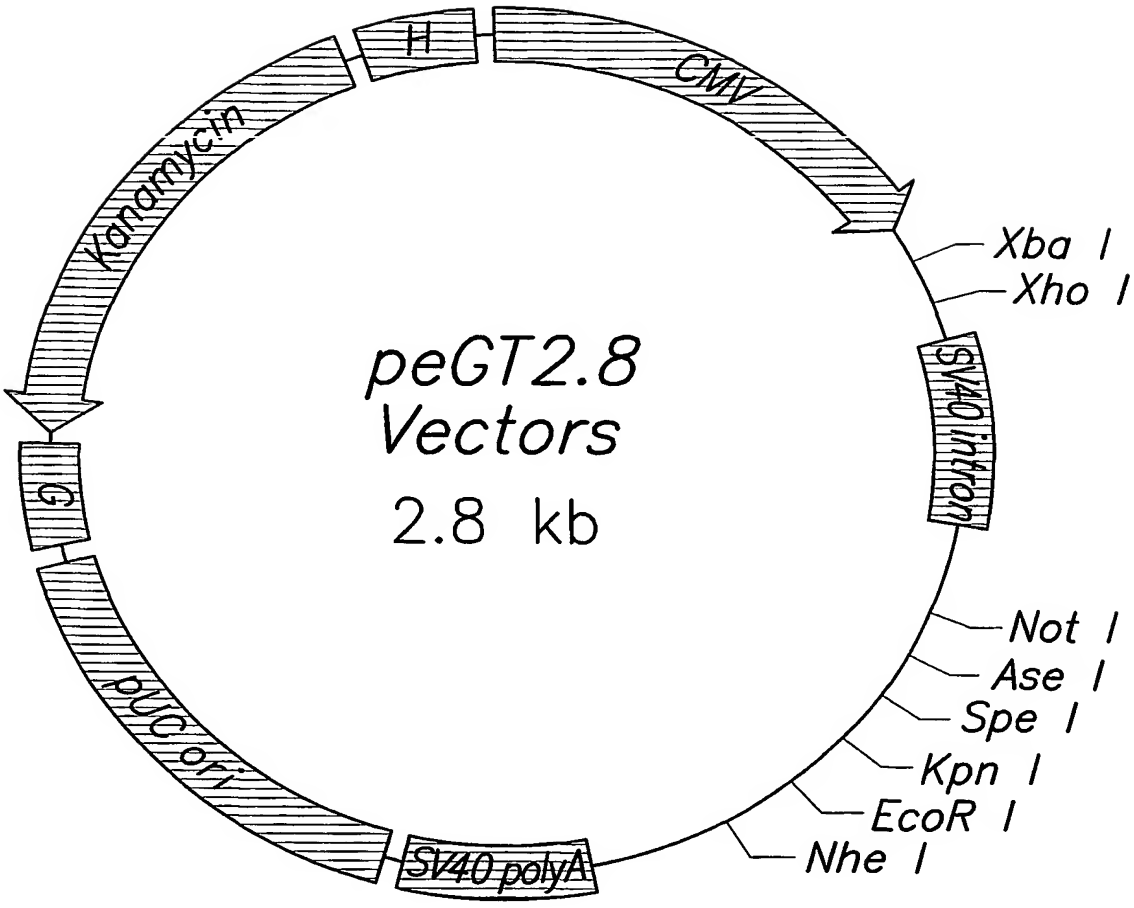
tggatgaaga ccagcccttc ccggctgtgc cgaaatggtc catcaaaaaa tggctttcgc 2520
tacctggaga gacgcgcccg ctgatccttt gcgaatacgc ccacgcgatg ggtaacagtc 2580
ttggcggttt cgctaaatac tggcaggcgt ttcgtcagta tccccgttta cagggcggct 2640
tcgtctggga ctgggtggat cagtcgctga ttaaatatga tgaaaacggc aacccgtggt 2700
cggcttacgg cggtgatttt ggcgatacgc cgaacgatcg ccagttctgt atgaacggtc 2760
tggtctttgc cgaccgcacg ccgcatccag cgctgacgga agcaaaacac cagcagcagt 2820
ttttccagtt ccgtttatcc gggcaaacca tcgaagtgac cagcgaatac ctgttccgtc 2880
atagcgataa cgagctcctg cactggatgg tggcgctgga tggtaagccg ctggcaagcg 2940
gtgaagtgcc tctggatgtc gctccacaag gtaaacagtt gattgaactg cctgaactac 3000
cgcagccgga gagcgccggg caactctggc tcacagtacg cgtagtgcaa ccgaacgcga 3060
ccgcatggtc agaagccggg cacatcagcg cctggcagca gtggcgtctg gcggaaaacc 3120
tcagtgtgac gctccccgcc gcgtcccacg ccatcccgca tctgaccacc agcgaaatgg 3180
atttttgcat cgagctgggt aataagcgtt ggcaatttaa ccgccagtca ggctttcttt 3240
cacagatgtg gattggcgat aaaaaacaac tgctgacgcc gctgcgcgat cagttcaccc 3300
gtgcaccgct ggataacgac attggcgtaa gtgaagcgac ccgcattgac cctaacgcct 3360
gggtcgaacg ctggaaggcg gcgggccatt accaggccga agcagcgttg ttgcagtgca 3420
cggcagatac acttgctgat gcggtgctga ttacgaccgc tcacgcgtgg cagcatcagg 3480
ggaaaacctt atttatcagc cggaaaacct accggattga tggtagtggt caaatggcga 3540
ttaccgttga tgttgaagtg gcgagcgata caccgcatcc ggcgcggatt ggcctgaact 3600
gccagctggc gcaggtagca gagcgggtaa actggctcgg attagggccg caagaaaact 3660
atcccgaccg ccttactgcc gcctgttttg accgctggga tctgccattg tcagacatgt 3720
ataccccgta cgtcttcccg agcgaaaacg gtctgcgctg cgggacgcgc gaattgaatt 3780
atggcccaca ccagtggcgc ggcgacttcc agttcaacat cagccgctac agtcaacagc 3840
aactgatgga aaccagccat cgccatctgc tgcacgcgga agaaggcaca tggctgaata 3900
tcgacggttt ccatatgggg attggtggcg acgactcctg gagcccgtca gtatcggcgg 3960
aattacagct gagcgccggt cgctaccatt accagttggt ctggtgtcaa aaataataat 4020
aaccgggcag gccatgtctg cccgtatttc gcgtaaggaa atccattatg tactatttaa 4080
aaaacacaaa cttttggatg ttcggtttat tctttttctt ttacttttttt atcatgggag 4140
cctacttccc gtttttcccg atttggctac atgacatcaa ccatatcagc aaaagtgata 4200
cgggtattat ttttgccgct atttctctgt tctcgctatt attccaaccg ctgtttggtc 4260
tgctttctga caaactcggc ctcgactcta ggcggccgcg gggatccaga catgataaga 4320
tacattgatg agtttggaca aaccacaact agaatgcagt gaaaaaaatg ctttatttgt 4380
gaaatttgtg atgctattgc tttatttgta accattataa gctgcaataa acaagttaac 4440
aacaacaatt gcattcattt tatgtttcag gttcaggggg aggtgtggga ggtttttttcg 4500
gatcctctag agtcgacctg caggcatgca agcttggcgt aatcatggtc atagctgttt 4560
cctgtgtgaa attgttatcc gctcacaatt ccacacaaca tacgagccgg aagcataaag 4620
tgtaaagcct ggggtgccta atgagtgagc taactcacat taattgcgtt gcgctcactg 4680
cccgctttcc agtcgggaaa cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg 4740
gggagaggcg gtttgcgtat tgggcgctct tccgcttcct cgctcactga ctcgctgcgc 4800
tcggtcgttc ggctgcggcg agcggtatca gctcactcaa aggcggtaat acggttatcc 4860
acagaatcag gggataacgc aggaaagaac atgtgagcaa aaggccagca aaaggccagg 4920
aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc tccgcccccc tgacgagcat 4980
cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga caggactata aagataccag 5040
gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga 5100
tacctgtccg cctttctccc ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg 5160
tatctcagtt cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga acccccgtt 5220
cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc ggtaagacac 5280
gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc 5340

ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag gacagtattt 5400
ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc 5460
ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt gcaagcagca gattacgcgc 5520
agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga cgctcagtgg 5580
aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat cttcacctag 5640
atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga gtaaacttgg 5700
tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg tctatttcgt 5760
tcatccatag ttgcctgact ccccgtcgtg tagataacta cgatacggga gggcttacca 5820
tctggcccca gtgctgcaat gataccgcga gacccacgct caccggctcc agatttatca 5880
gcaataaacc agccagccgg aagggccgag cgcagaagtg gtcctgcaac tttatccgcc 5940
tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc agttaatagt 6000
ttgcgcaacg ttgttgccat tgctacaggc atcgtggtgt cacgctcgtc gtttggtatg 6060
gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc catgttgtgc 6120
aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca gaagtaagtt ggccgcagtg 6180
ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc atccgtaaga 6240
tgcttttctg tgactggtga gtactcaacc aagtcattct gagaatagtg tatgcggcga 6300
ccgagttgct cttgcccggc gtcaatacgg gataataccg cgccacatag cagaacttta 6360
aaagtgctca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat cttaccgctg 6420
ttgagatcca gttcgatgta acccactcgt gcacccaact gatcttcagc atcttttact 6480
ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata 6540
agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta ttgaagcatt 6600
tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacaa 6660
ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt 6720
atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtct cgcgcgtttc 6780
ggtgatgacg gtgaaaacct ctgacacatg cagctcccgg agacggtcac agcttgtctg 6840
taagcggatg ccgggagcag acaagcccgt cagggcgcgt cagcgggtgt tggcgggtgt 6900
cggggctggc ttaactatgc ggcatcagag cagattgtac tgagagtgca ccatatgcgg 6960
tgtgaaatac cgcacagatg cgtaaggaga aaataccgca tcaggcgcca ttcgccattc 7020
aggctgcgca actgttggga agggcgatcg gtgcgggcct cttcgctatt acgccagctg 7080
gcgaaagggg gatgtgctgc aaggcgatta agttgggtaa cgccagggtt ttcccagtca 7140
cgacgttgta aaacgacggc cagt 7164



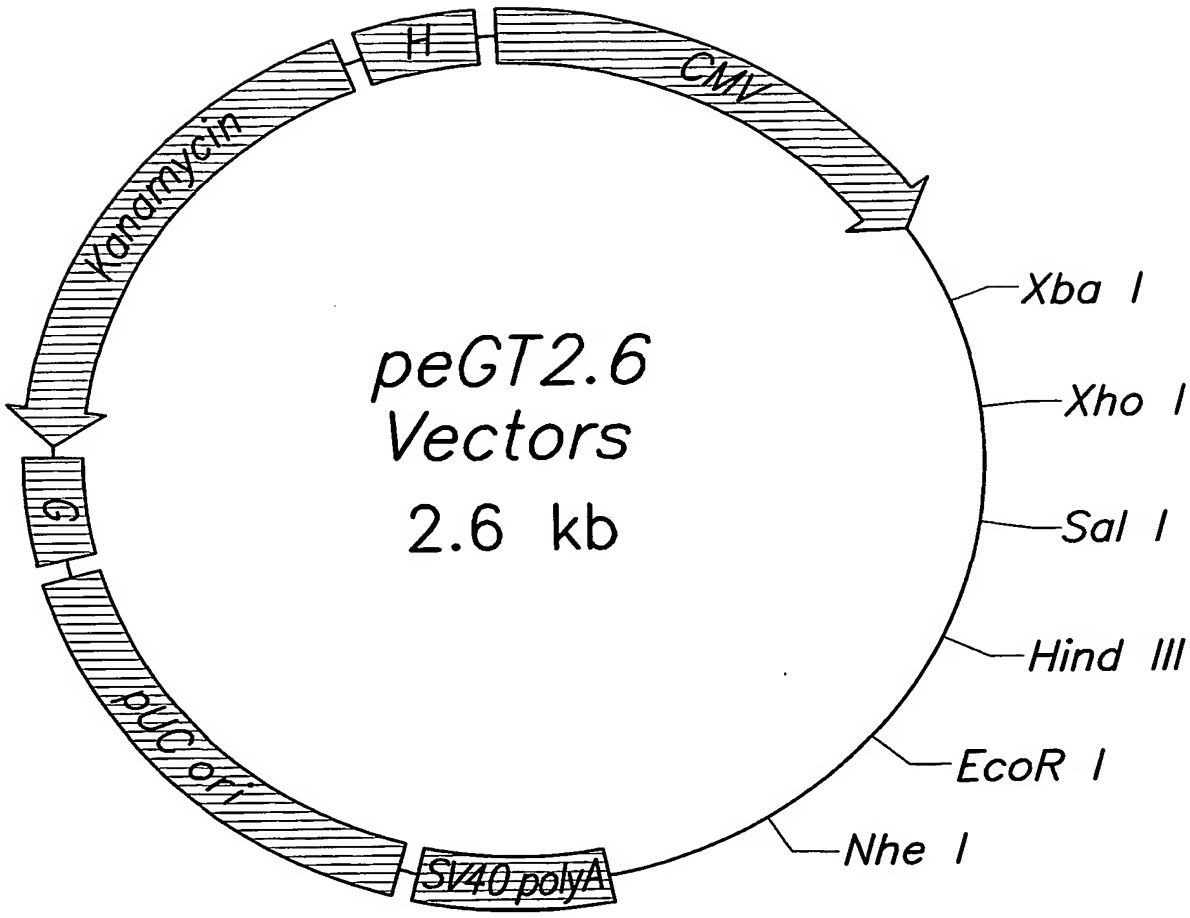
peGT β: G=H=0
peGT β –ISS2: G=H=1 ISS
peGT β –ISS4: G=H=2 ISS

FIG. 1



peGT2.8: G=H=0
peGT2.8–ISS2: G=H=1 ISS
peGT2.8–ISS4: G=H=2 ISS

FIG. 2



peGT2.6: G=H=0
peGT2.6–ISS2: G=H=1 ISS
peGT2.6–ISS4: G=H=2 ISS

FIG. 3

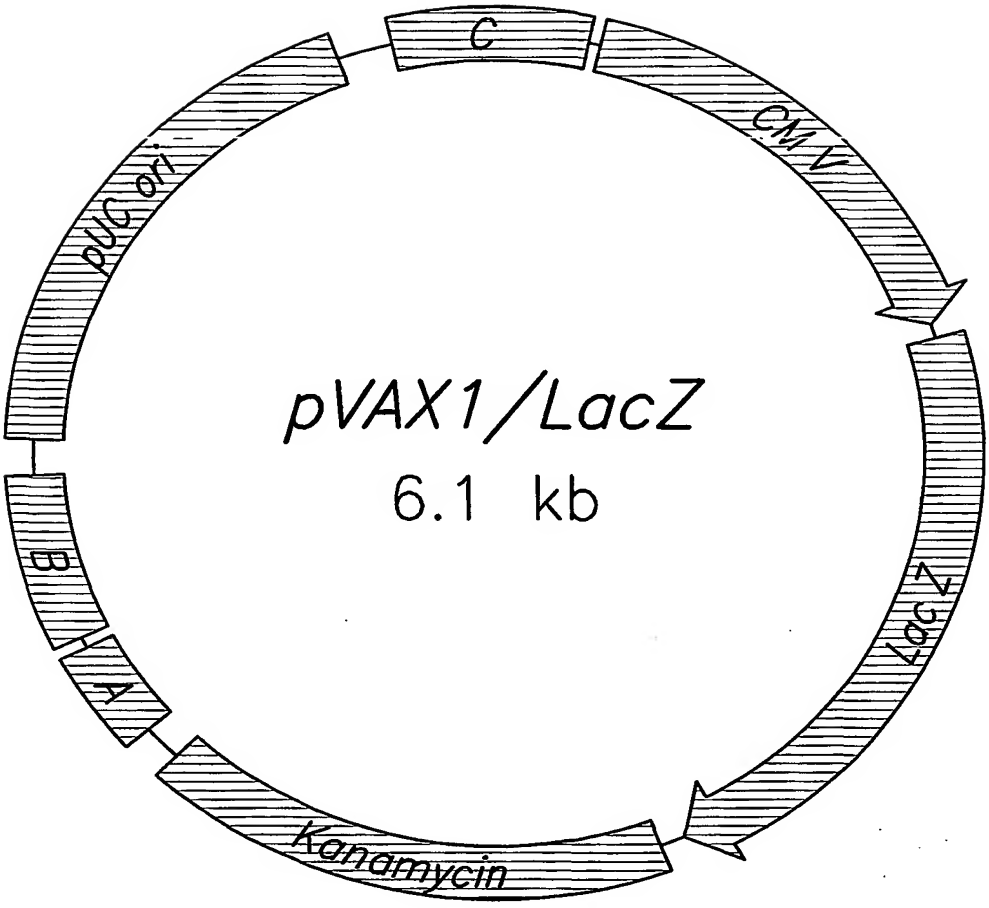


FIG. 4

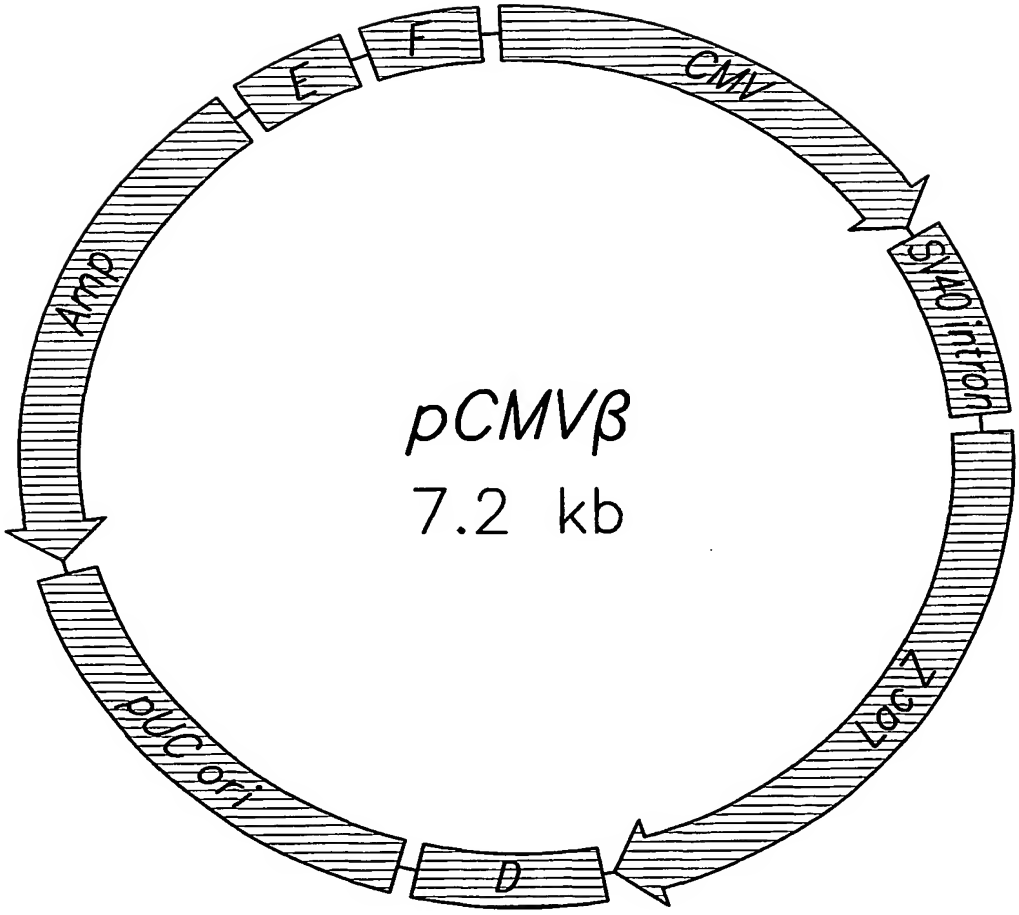


FIG. 5

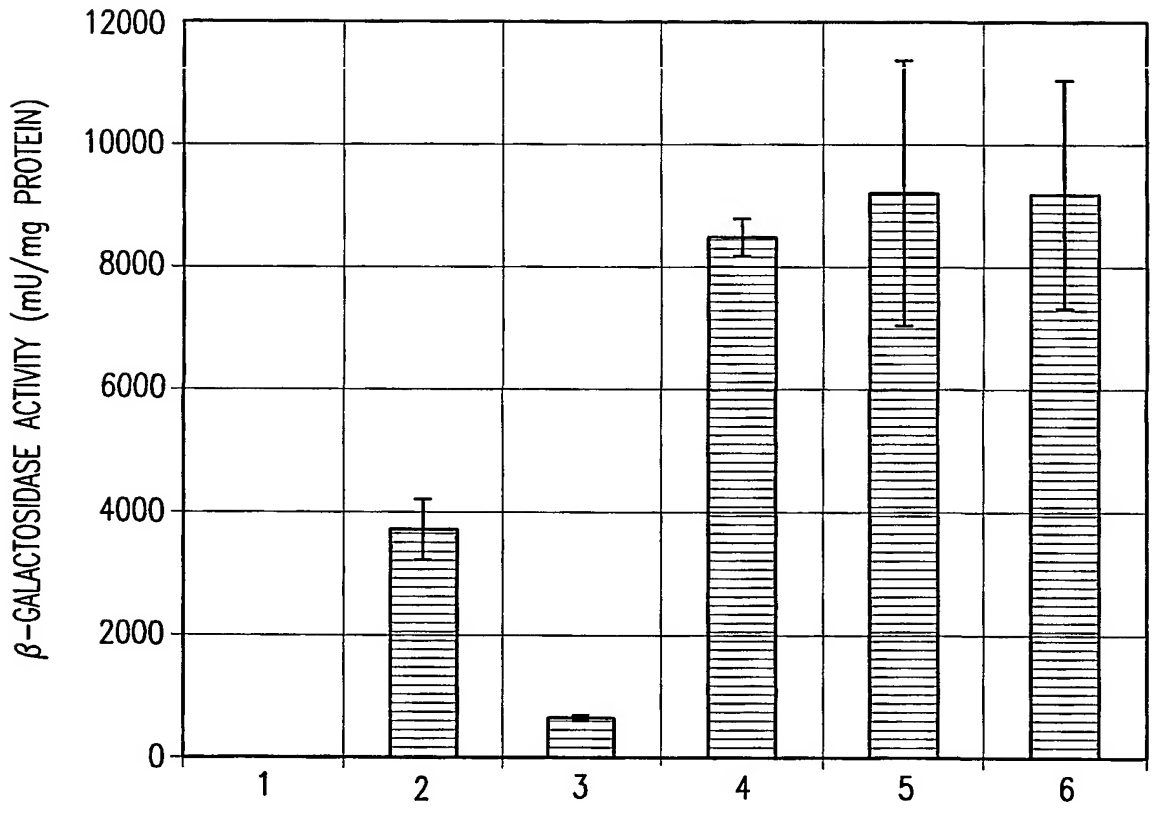


FIG. 6

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 July 2001 (19.07.2001)

PCT

(10) International Publication Number
WO 01/51626 A3

(51) International Patent Classification[7]: C12N 15/85, 15/67, A61K 48/00

(21) International Application Number: PCT/US01/01255

(22) International Filing Date: 9 January 2001 (09.01.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/480,879 10 January 2000 (10.01.2000) US

(71) Applicant: ELIM BIOPHARMACEUTICALS, INC. [US/US]; 901 Grandview Drive, South San Francisco, CA 94080 (US).

(72) Inventors: LU, Xianbin; 311 Faxon Avenue, San Francisco, CA 94112 (US). SUN, Lihyun; 140 Ware Road, Woodside, CA 94062 (US). ZHANG, Yilin; 140 Ware Road, Woodside, CA 94062 (US).

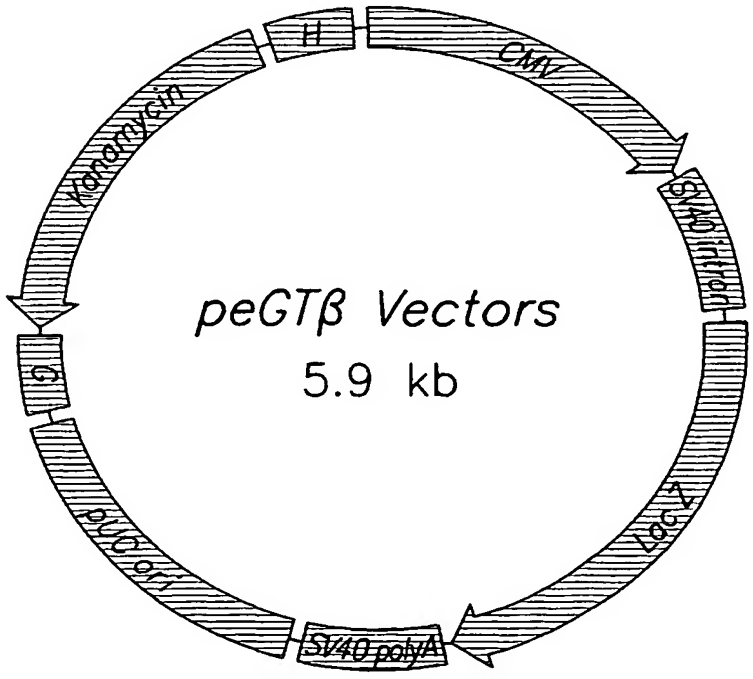
(74) Agents: CHOW, Y., Ping et al.; Heller Ehrman White & McAuliffe LLP, 275 Middlefield Road, Menlo Park, CA 94025-3506 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

[Continued on next page]

(54) Title: NOVEL PLASMID EXPRESSION VECTORS

(57) Abstract: This invention relates to plasmid backbones and constructs wherein non-essential nucleotide sequences have been substantially removed. These backbones can be used to clone genes of interest. The constructs are useful in gene therapy, DNA vaccines and the *in vitro* production of polypeptides and/or proteins.



peGT β: G=H=0
peGT β -ISS2: G=H=1 ISS
peGT β -ISS4: G=H=2 ISS

WO 01/51626 A3

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report*

(88) **Date of publication of the international search report:**
14 February 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

Inte tional Application No
PC1/US 01/01255

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/85 C12N15/67 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HARTIKKA J: "AN IMPROVED PLASMID DNA EXPRESSION VECTOR FOR DIRECT INJECTION INTO SKELETAL MUSCLE" HUMAN GENE THERAPY, vol. 7, no. 10, 20 June 1996 (1996-06-20), pages 1205-1217, XP002050079 ISSN: 1043-0342	1-14
Y	page 1206, left-hand column; table 1 page 1210, right-hand column -page 1211, left-hand column, paragraph 2	15

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
6 July 2001	31/07/2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Steffen, P

2

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	NORMAN JON A ET AL: "Development of improved vectors for DNA-based immunization and other gene therapy applications." VACCINE, vol. 15, no. 8, 1997, pages 801-803, XP004075655 ISSN: 0264-410X page 801, left-hand column -page 802, right-hand column, paragraph 1; figure 1	1-14
Y	abstract	15
X	MONTGOMERY DONNA L ET AL: "Heterologous and homologous protection against influenza A by DNA vaccination: Optimization of DNA vectors." DNA AND CELL BIOLOGY, vol. 12, no. 9, 1993, pages 777-783, XP000565708 ISSN: 1044-5498	1-13
Y	page 779, left-hand column -page 780, left-hand column, paragraph 1; figure 1	14,15
X	INVITROGEN: "pVAX1 (Catalog no. V260-20) Version B 010124/25-0256" INVITROGEN HOMEPAGE (WWW.INVITROGEN.COM), 'Online! 1998 - 2001, XP002171177 Retrieved from the Internet: <URL:http://www.invitrogen.com/Content/Tech-Online/molecular_biology/manuals_pps/pvax1_man.pdf> 'retrieved on 2001-07-03!	1-6, 11-13
Y	page 1 page 5 -page 7	7-10,14, 15
X	LANG C ET AL: "Efficient expression and secretion of Aspergillus niger RH5344 polygalacturonase in Saccharomyces cerevisiae." APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, vol. 44, no. 1-2, 1995, pages 147-156, XP001005274 ISSN: 0175-7598 page 150, left-hand column, paragraph 3 -page 151, right-hand column, paragraph 3; figure 2; table 1	1-5, 11-13

-/--

2

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	AZEVEDO V ET AL: "Main features of DNA-based immunization vectors." BRAZILIAN JOURNAL OF MEDICAL AND BIOLOGICAL RESEARCH, vol. 32, no. 2, February 1999 (1999-02), pages 147-153, XP002171178 ISSN: 0100-879X page 149 -page 152, right-hand column, paragraph 1; figure 1 ---	1-15
Y	HASAN U A ET AL: "Nucleic acid immunization: Concepts and techniques associated with third generation vaccines." JOURNAL OF IMMUNOLOGICAL METHODS, vol. 229, no. 1-2, 29 October 1999 (1999-10-29), pages 1-22, XP001000923 ISSN: 0022-1759 page 3, right-hand column page 7 ---	1-15
Y	KAY M A ET AL: "GENE THERAPY" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA,US,NATIONAL ACADEMY OF SCIENCE. WASHINGTON, vol. 94, November 1997 (1997-11), pages 12744-12746, XP002922034 ISSN: 0027-8424 page 12744, left-hand column, paragraph 3 page 12476, left-hand column ---	14,15
A	GIESE MATTHIAS: "DNA-antiviral vaccines: New developments and approaches: A review." VIRUS GENES, vol. 17, no. 3, 1998, pages 219-232, XP001005248 ISSN: 0920-8569 page 220, right-hand column, paragraph 3 -page 221, right-hand column, paragraph 1 --- -/--	

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	CLONTECH: "Vector Info pCMVbeta" CLONTECH HOMEPAGE (WWW.CLONTECH.COM), ‘Online! 1997, XP002171179 Retrieved from the Internet: <URL:http://www.clontech.com/techinfo/vectors/vectorsC/pCMVbeta.shtml#description> ‘retrieved on 2001-07-03! the whole document	
A	-& CLONTECH: "pCMVbeta Vector Information" CLONTECH HOMEPAGE (WWW.CLONTECH.COM), ‘Online! 1997 - 1997, XP002171180 Retrieved from the Internet: <URL:<URL:http://www.clontech.com/techinfo/vectors/vectorsC/pdf/PT2004-5.pdf>> ‘retrieved on 2001-07-03! the whole document	
P,X	WO 00 40737 A (KIM DUK KYUNG ;LEE YOUNG JOO (KR); YU SEUNG SHIN (KR); KIM SUNYOUN) 13 July 2000 (2000-07-13) page 19, line 12 -page 21, line 20; figures 1,4-6	1-15

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/01255

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0040737 A	13-07-2000	AU 1896500 A	24-07-2000